

*Review*

# Genomics-Assisted Cassava Breeding: A Comprehensive Review of Advances, Tools, and Future Prospects

**Simon Peter Abah \*, Mbe Joseph Okpani, Ihuoma Umezurumba Okwuonu and Chiedozie Ngozi Egesi**

National Root Crops Research Institute, Umudike, Umuahia P.M.B. 7006, Abia State, Nigeria; mbejosepho@gmail.com (MJO); ihuomaumezurumba@gmail.com (IUA); cegesi@yahoo.com (CNE)

\* Correspondence: Simon Peter Abah, Email: abahsp@gmail.com

---

## ABSTRACT

Cassava (*Manihot esculenta* Crantz) is a critical tropical root crop that underpins food security, industrial applications, and rural livelihoods, particularly across sub-Saharan Africa. Traditional cassava breeding has faced significant challenges due to the crop's long growth cycle, clonal propagation, high heterozygosity, polygenic inheritance of key traits, and strong genotype  $\times$  environment (G  $\times$  E) interactions. These factors have historically limited the rate of genetic improvement, despite cassava's socio-economic importance. Recent developments in genomics-assisted breeding (GAB) are transforming cassava improvement by combining high-density molecular markers, genome-wide association studies (GWAS), genomic selection (GS), and multi-omics approaches such as transcriptomics, metabolomics, and high-throughput phenotyping. The sequencing of the cassava reference genome has enabled single nucleotide polymorphism (SNP) discovery, quantitative trait loci (QTL) mapping, and the development of genomic prediction models, allowing early selection for complex traits including root yield, dry matter content, starch quality, carotenoid accumulation, disease resistance, and early storage root bulking. Functional genomics, together with non-invasive root imaging technologies such as ground-penetrating radar, provides insight into the genetic and physiological mechanisms underlying root development and stress responses. Multi-environment GWAS and genomic selection models now account for G  $\times$  E interactions, improving prediction accuracy and supporting rapid-cycle breeding. Furthermore, integrating metabolomic and transcriptomic datasets enhances trait prediction, accelerates candidate gene discovery, and advances precision breeding strategies. This review provides recent progress, tools, and future prospects in genomics-assisted cassava breeding, illustrating how modern genomic and phenomic technologies are reshaping breeding pipelines, shortening breeding cycles, and facilitating the development of climate-resilient, high-yielding, and farmer-preferred cassava varieties.

## Open Access

Received: 3 Dec 2025

Accepted: 22 Jan 2026

Published: 28 Jan 2026

Copyright © 2026 by the author. Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of Creative Commons Attribution 4.0 International License.

**KEYWORDS:** genomics-assisted breeding; cassava (*Manihot esculenta*); genomic selection; multi-omics integration; trait prediction and precision breeding

## INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is a staple food and industrial crop of global importance, providing dietary energy for over 800 million people and supporting rural livelihoods across sub-Saharan Africa, Latin America, and Southeast Asia [1,2]. Its ability to produce reasonable yields under marginal soils, intermittent drought, and low external inputs has made cassava a strategic crop for food security in the face of climate variability [3,4]. In Africa, particularly Nigeria, the world's largest producer, cassava underpins household food systems, starch-based industries, and emerging bioeconomy value chains [5,6].

Despite its resilience and socio-economic relevance, genetic improvement in cassava has lagged behind that of major cereal crops. This slow progress is largely attributable to inherent biological and genetic constraints, including a long breeding cycle (8–18 months), clonal propagation, high heterozygosity, and severe inbreeding depression [7–9]. Moreover, most agronomically and industrially important traits in cassava, such as storage root yield, dry matter content, starch quality, carotenoid accumulation, and cyanogenic potential, are quantitatively inherited and strongly influenced by genotype  $\times$  environment (G  $\times$  E) interactions [10,11]. These features reduce selection efficiency under conventional phenotypic breeding and necessitate long, resource-intensive multi-location testing schemes [12].

Over the past two decades, cassava breeding has undergone a major transformation driven by advances in genomics, particularly the release of high-quality reference genome assemblies and the development of cost-effective, high-density single nucleotide polymorphism (SNP) genotyping platforms [13–15]. These genomic resources, combined with declining genotyping costs, have enabled the systematic application of genomics-assisted breeding (GAB) approaches in cassava, including marker-assisted selection (MAS), genomic selection (GS), genome-wide association studies (GWAS), and more recently, genome editing using CRISPR/Cas systems [11,16–19].

Marker-assisted selection has been successfully deployed in cassava breeding programs, particularly for traits controlled by major-effect loci such as resistance to cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) [9,11,16,17,20]. However, its effectiveness diminishes for complex quantitative traits, including storage root yield, dry matter content, and starch quality, which are governed by many small-effect loci and are highly influenced by genotype  $\times$  environment interactions [10,11]. These limitations have motivated a shift toward genomic selection, which leverages genome-wide marker information to

capture cumulative additive genetic effects without requiring prior knowledge of individual quantitative trait loci [21,22].

Empirical studies have demonstrated the potential of genomic selection to improve prediction accuracy and reduce breeding cycle length in cassava, although reported accuracies vary widely across traits, populations, and environments [23]. Variability in GS performance has been attributed to differences in training population size, genetic relatedness between training and prediction sets, marker density, phenotyping precision, and environmental heterogeneity [22,24]. These findings highlight the critical role of high-quality phenotyping and multi-environment data in realizing the full benefits of genomic prediction in cassava breeding.

In parallel, GWAS has contributed to the identification of genomic regions and candidate genes associated with key agronomic and quality traits, including root architecture, carotenoid accumulation, cyanogenic potential, and disease resistance [16,17,25,26]. While GWAS has enhanced biological understanding of trait architecture, the translation of association signals into routine breeding decisions remains challenging due to small effect sizes and limited reproducibility across environments.

More recently, genome editing technologies, particularly CRISPR/Cas-based systems, have emerged as powerful tools for functional genomics and targeted trait improvement in cassava. Successful edits have been reported for genes involved in disease resistance, starch biosynthesis, and cyanogenic glucoside metabolism [27,28]. Nevertheless, the widespread adoption of genome editing in cassava breeding is constrained by low transformation efficiency, genotype dependency, regulatory uncertainty in many cassava-growing regions, and incomplete functional annotation of the cassava genome [19,29].

This review is therefore guided by the central hypothesis that integrating genomic tools with high-quality phenotyping and environmental information can substantially enhance selection accuracy and accelerate genetic gain in cassava breeding, particularly for complex, polygenic traits. By examining empirical evidence across MAS, GS, GWAS, and genome editing studies, this review evaluates how genomics-assisted breeding has reshaped cassava improvement efforts, identifies critical bottlenecks limiting its full adoption, and highlights emerging opportunities for next-generation breeding frameworks.

## Genomic-Assisted Breeding as a Catalyst for Accelerated Improvement

The advent of molecular markers, cost-effective sequencing, and high-resolution phenotyping has fundamentally transformed cassava improvement over the past decade. Genomics-assisted breeding (GAB) integrates next-generation sequencing (NGS), genome-wide association studies (GWAS), genomic selection (GS), and multi-omics data

(transcriptomics, metabolomics, and phenomics) to unravel trait architecture and accelerate breeding progress [11,30,31]. These innovations follow the sequencing of the cassava reference genome [32], which laid the foundation for large-scale SNP discovery, high-density SNP arrays, and genotyping-by-sequencing platforms [33].

Genomic prediction tools developed through international consortia, such as the NextGen Cassava Breeding Project, enable breeders to estimate genomic estimated breeding values (GEBVs) early in the selection cycle, bypassing lengthy field evaluations and enhancing response to selection [11,34]. Combined with high-throughput phenotyping technologies, including unmanned aerial systems, imaging platforms, and non-invasive root imaging methods such as ground penetrating radar [35], genomic tools allow quantification of complex traits that previously confounded traditional breeding.

## CASSAVA BIOLOGY AND GENETIC IMPROVEMENT CONSTRAINTS

### Plant Biology and Reproductive System

Cassava (*Manihot esculenta* Crantz) exhibits a predominantly outcrossing reproductive system, a characteristic that has profound implications for its genetic structure and breeding behavior. Although a small degree of natural selfing does occur, estimated at 5%–10%, outcrossing is maintained largely through asynchronous flowering, spatial arrangement of male and female flowers on the same inflorescence, and innate self-incompatibility tendencies [9,36,37].

From a population genetics perspective, this mating system contributes to:

**Extreme heterozygosity** which means cassava cultivars carry numerous heterozygous loci accumulated through generations of vegetative propagation and outcrossing. This maintains broad genetic diversity within clones but complicates allele fixation.

**High segregation in progeny** which means sexual recombination results in marked phenotypic variability, even among full-sib families, often producing many inferior segregants and only a few elite recombinants [38].

**Severe inbreeding depression** which means attempts to self-pollinate cassava drastically reduces vigor, fertility, root yield, and even survival rates. This is attributed to the expression of deleterious recessive alleles that accumulate over clonal generations [39].

**Difficulty generating homozygous lines**, which means traditional line development, as practiced in cereals, is nearly impossible in cassava. The few cycles of selfing achievable often yield weak plants, thereby slowing the development of inbred lines needed for hybrid breeding [36].

These biological limitations fundamentally shape cassava breeding strategies, forcing reliance on clonal selection, recurrent phenotypic evaluation, and more recently, genomic tools capable of capturing additive and non-additive genetic variance in highly heterozygous populations.

### Storage Root Formation

Cassava storage root development is a complex physiological process integrating anatomical differentiation, secondary growth, carbon assimilation, and hormonal regulation. Unlike tuber crops such as potato, cassava storage roots originate from fibrous roots that undergo secondary thickening. Key developmental processes include:

**Phloem parenchyma cell differentiation** which involves the enlargement and starch deposition that occur primarily in the secondary phloem. The early differentiation of parenchyma cells sets the foundation for bulking capacity [40].

**Secondary growth** which involves the cambial activity, producing secondary xylem and secondary phloem, that drives root thickening. Varieties differ widely in how early and vigorously these tissues proliferate [41].

**Carbohydrate partitioning** which involves the allocation of photosynthates between shoots biomass and root sinks that determines early bulking. Genotypes with stronger sink strength and higher assimilate translocation tend to bulk earlier [42].

**Hormonal control** which involves the auxins that regulate cambial activity and root thickening, cytokinins promote cell division, abscisic acid (ABA) is associated with stress-induced root growth, and gibberellins influence starch accumulation. The interplay of these hormones strongly affects bulking patterns [43].

Importantly, early storage root bulking is highly genotype-dependent. Rüscher et al. (2021) [41] demonstrated that cultivar-specific differences in cambial activation, sugar transporter expression, and phloem/xylem ratios largely dictate how early storage roots develop. Environmental factors, especially soil moisture, temperature, and nutrient availability, may enhance or suppress bulking, but the genetic blueprint remains the primary determinant.

### Key Breeding Challenges

Cassava breeders face several interconnected constraints, many of which stem directly from the crop's biology and the complexity of target traits (Table 1). These bottlenecks slow genetic gain and limit the pace of varietal turnover.

**Table 1.** Key Breeding Challenges and their implication for breeding.

Challenge	Implication for Breeding
Long growth cycle (8–18 months)	Reduces selection turnover, increases cost of trials, and slows genetic gain per year [36].
Vegetative propagation	Enhances clonal fidelity but increases risk of accumulating systemic diseases such as CMD, CBSD, and bacterial blight. Limits distribution of clean planting materials [44].
Polygenic traits (yield, starch content, root size)	Require numerous markers for accurate estimation of breeding values; MAS is insufficient, making GS and GWAS indispensable [11].
Strong G×E interactions	Trait performance varies widely across environments; reduces selection accuracy and complicates multi-location testing [45].
Limited phenotyping capacity	Many traits, especially root traits, are difficult to measure reliably, resulting in poor-quality data that weaken genomic predictions [35].

## GENETIC RESOURCES AND DIVERSITY IN CASSAVA

The foundation of any successful crop-improvement program lies in the depth and breadth of available genetic resources. Cassava benefits from one of the most geographically and genetically diverse germplasm pools among tropical root crops, owing to its long history of domestication in the Amazon Basin and subsequent diffusion across Africa and Asia. The crop was introduced to Africa in the 16th century and subsequently adapted to a wide range of agro-ecologies, producing thousands of distinct landraces with unique quality, resistance, and culinary properties [46,47].

Today, the conservation, characterization, and utilization of cassava genetic diversity are anchored within several global genebanks and national breeding programs. These collections safeguard not only traditional landraces but also elite breeding lines, wild relatives, and improved varieties that serve as essential materials for breeding, genomic analyses, and prebreeding research.

### Major Cassava Global Collections

International Center for Tropical Agriculture (CIAT), Colombia has ~6,000 accessions as the largest international cassava collection, representing broad genetic diversity from South America, the center of origin and diversification. CIAT's genebank includes wild *Manihot* species, indigenous landraces, and elite breeding materials, making it a core resource for global prebreeding and genomics research [48].

International Institute of Tropical Agriculture (IITA), Nigeria maintains the most extensive African cassava collection, comprising landraces, farmer-preferred clones, and advanced breeding lines with resistance to cassava mosaic disease (CMD) and cassava brown streak disease (CBSD). IITA materials have been foundational to African varietal improvement and genomic selection initiatives [30].

National Root Crops Research Institute (NRCRI), Nigeria as one of the major custodians of Nigerian cassava diversity, NRCRI holds landraces associated with culturally specific products such as fufu, gari, abacha, and lafun. These accessions are invaluable for improving culinary quality, textural properties, fermentation behavior, and processing efficiency, traits that are increasingly important for consumer- and industry-oriented breeding.

Brazilian Agricultural Research Corporation (EMBRAPA), Brazil houses a vast collection of South American germplasm, including wild relatives, bitter and sweet cassava types, and industrial varieties with diverse starch profiles. These accessions represent the deepest reservoir of allelic diversity for traits such as root architecture, drought tolerance, and starch biochemistry [49].

Together, these collections ensure that cassava breeding programs have access to wide genetic variation essential for addressing emerging challenges, including climate stress, evolving pathogens, changing consumer preferences, and industrial processing demands.

### Molecular Diversity Studies

Advances in molecular markers, particularly SNPs, have transformed the characterization of cassava genetic resources. High-density SNP datasets generated through GBS and SNP arrays have provided new insights into population structure, genetic relationships, and trait-associated alleles across global collections.

Key findings from molecular diversity studies reveal that:

Strong population differentiation exists between African and Latin American germplasm. Because cassava originated in South America, African populations represent a subset of global diversity but have undergone extensive local adaptation. SNP-based analyses consistently show clear genetic partitioning between these regions [23,50].

Deep allelic diversity is found for CMD resistance, carotenoid content, and starch physicochemical properties with several major-effect loci for disease resistance, such as the CMD2 locus, have been identified mainly in African germplasm, while Latin American accessions often harbor unique alleles influencing carotenoids, root dry matter, branching patterns, and cyanogenic potential [17].

Low linkage disequilibrium (LD) across the genome enhances suitability for GWAS; LD in cassava decays rapidly due to outcrossing and high heterozygosity, allowing fine mapping of traits and high resolution in GWAS analyses [50,51]. This property makes cassava one of the more amenable clonally propagated crops for high-resolution genomic dissection.

These molecular insights have accelerated the discovery of markers for quality attributes, yield components, and stress tolerance, ultimately strengthening genomic-assisted breeding pipelines worldwide.

### Prebreeding Opportunities

Prebreeding, introgression of novel alleles from unadapted or wild germplasm into elite breeding pools, is increasingly recognized as a strategic pathway for expanding cassava's genetic base. Wild *Manihot* species offer reservoirs of resistance genes, stress-tolerance mechanisms, and value-added biochemical traits not present in cultivated accessions.

#### ***Wild Relatives as Sources of Novel Diversity***

*Manihot flabellifolia* which is widely regarded as one of the closest wild progenitors of cultivated cassava provides alleles for drought tolerance, vigor, and root biomass accumulation. Genetic studies have shown shared ancestry and potential for hybridization with cassava, making it a high-priority species for introgression [47].

*Manihot pruinosa* and *Manihot esculenta* ssp. *Peruviana* are species that harbor genes related to pest and disease resistance, including tolerance to bacterial blight and root rot. They also possess unique starch traits and root morphologies that may be useful for industrial applications such as high-viscosity starch, bioethanol, and biodegradable polymers [52].

#### ***Opportunities for Trait Enhancement***

Prebreeding with wild relatives offers opportunities to introduce: resistance to emerging diseases; improved drought and heat tolerance; novel starch composition and granule morphology; altered cyanogenic glucoside content; enhanced root structural properties for early bulking and mechanization; unique secondary metabolites for biofortification or industrial uses [7,11,14,39].

However, introgression remains challenging due to reproductive barriers, linkage drag, and the need for multiple cycles of backcrossing and genomic selection. Modern tools such as marker-assisted backcrossing, genomic selection, and genome editing may shorten prebreeding cycles and help capture beneficial alleles while minimizing undesirable traits.

## **MOLECULAR MARKER TECHNOLOGIES IN CASSAVA BREEDING**

Molecular marker technologies have played a central role in transforming cassava breeding from a largely phenotype-driven enterprise into a modern, genomics-assisted discipline. These tools have enabled deeper understanding of the genetic architecture underlying key agronomic, quality, and adaptive traits, while providing the foundation for MAS, GWAS, and GS. The evolution from low-throughput, anonymous

markers to high-density, sequence-based platforms have significantly accelerated genetic gain in cassava improvement programs worldwide.

### Evolution of Marker Systems

The application of molecular markers in cassava has evolved over three decades, mirroring advancements in global plant genomics. Each generation of markers addressed specific limitations of its predecessor, ultimately culminating in genome-wide SNP platforms that now underpin modern cassava breeding pipelines.

#### ***Early Marker Systems: RAPD and AFLP***

Random Amplified Polymorphic DNA (RAPD) and Amplified Fragment Length Polymorphism (AFLP) markers were among the earliest tools used to examine cassava genetic diversity in the 1990s and early 2000s. These markers were relatively inexpensive and required limited prior sequence information, enabling rapid assessment of genetic relationships among landraces and wild *Manihot* species [53]. However, both systems suffered from limited reproducibility, dominance effects, and low marker density, factors that limited their utility for mapping complex traits.

#### ***Simple Sequence Repeats (SSRs): Towards Structured Genotyping***

The advent of SSR markers marked a major improvement in cassava molecular genetics. SSRs are co-dominant, highly polymorphic, and more reproducible than RAPD/AFLP markers. They enabled the construction of the first cassava linkage maps, provided the initial insights into population structure, and supported early QTL analyses for yield and disease resistance [54]. SSRs also facilitated germplasm fingerprinting and the management of global genebanks at CIAT, IITA, and EMBRAPA [54].

#### ***Single Nucleotide Polymorphisms (SNPs): The Genomic Revolution***

The sequencing of the cassava reference genome [32] marked a turning point in marker development. In the past decade, SNP markers have become the backbone of cassava genomic research and breeding because they offer: high marker density which enables fine-scale dissection of genetic variation; genome-wide coverage suitable for association mapping and genomic prediction; high reproducibility across laboratories and platforms; and scalability for routine use in breeding pipelines, including early-generation selection.

Several SNP platforms now widely support cassava breeding, such as:

**Genetic Enhancement of Cassava (GEM) SNP Chip**, developed under the NextGen Cassava project, offers thousands of carefully curated SNPs associated with key traits [11].

**Diversity Arrays Technology sequencing (DArTseq)** provides cost-effective, high-throughput SNP discovery for GWAS, genomic prediction, and diversity studies [51].

**Genotyping-by-Sequencing (GBS)** is a reduced-representation sequencing approach enabling generation of tens of thousands of genome-wide SNPs [55].

**Whole Genome Sequencing (WGS)-derived SNPs** is used for high-resolution mapping, evolutionary studies, and identification of structural variants [56–58].

Today, SNP-based genotyping is fully integrated into the breeding pipelines of CIAT, IITA, NRCRI, and EMBRAPA [11,14,16,19]. It forms the foundation of genomic selection, early-generation prediction, and rapid-cycle recurrent selection, strategies critical for increasing the rate of genetic gain in cassava.

## QTL Mapping Studies

Quantitative trait loci mapping has been instrumental in identifying genomic regions associated with disease resistance, quality traits, and yield components in cassava. Unlike major-gene crops, cassava improvement requires disentangling complex traits influenced by multiple loci and strong genotype  $\times$  environment interactions. Marker-based QTL studies provide the first step toward trait dissection, marker-assisted selection, and the design of genomic prediction models.

## Disease Resistance

Cassava diseases, particularly *cassava mosaic disease* and *cassava brown streak disease*, represent major constraints to yield stability and production across Africa. Molecular markers have facilitated significant breakthroughs in understanding and managing these diseases.

### Cassava Mosaic Disease

The CMD2 and CMD3 loci represent two of the most significant QTL discoveries in cassava disease resistance research.

**CMD2** is a dominant resistance locus originally identified in West African landraces. SNP markers tightly linked to CMD2 have allowed reliable marker-assisted selection in breeding programs [30].

**CMD3** confers quantitative resistance and complements CMD2 in breeding strategies targeting durable resistance.

The availability of molecular markers for these loci has greatly accelerated the release of CMD-resistant varieties throughout Africa.

### ***Cassava Brown Streak Disease***

QTL mapping and GWAS analyses have identified multiple genomic regions associated with CBSD tolerance, though resistance is more polygenic and environmentally sensitive than CMD. Markers flanking major-effect QTLs on chromosomes IV, VI, and XI have been reported in both CIAT and IITA populations [59]. Integrating these markers into genomic selection increases prediction accuracy for CBSD resistance under high-pressure environments.

### ***Quality Traits***

Cassava quality traits, including biofortification targets, processing qualities, and nutritional attributes, are increasingly central to breeding objectives. Molecular markers have shed light on the genetic architecture of these following traits in cassava:

Carotenoids (Pro-vitamin A) QTLs associated with  $\beta$ -carotene accumulation have been identified across multiple studies [60]. Markers linked to carotenoid biosynthesis genes such as *PSY2* and  *$\beta$ -carotene hydroxylase* support marker-assisted selection for yellow-fleshed biofortified varieties [61–63].

Several QTLs affecting dry matter, amylose:amylopectin ratio, granule size, and starch viscosity have been mapped, although most show moderate effect sizes. These traits are crucial for end-products such as gari, fufu, and industrial starch derivatives [64–66].

QTLs linked to cyanogenic glucosides on chromosomes II and XVI influence cyanide potential. Markers associated with *linamarin* and *lotaustralin* biosynthetic pathways help breeders balance safety, taste, and plant defense [67].

### ***Innovation in Root Trait Phenotyping and Early Bulking Improvement***

Early storage root bulking is a critical trait for market demand, drought escape, shortened production cycles, and farmer adoption across SSA [68,69]. However, the trait is quantitatively inherited, strongly influenced by the environment, and difficult to measure through destructive sampling. Recent advances in functional genomics and high-throughput phenotyping provide opportunities to dissect the genetic basis of early bulking and to model root system architecture *in situ*. Ground penetrating radar (GPR) provides moderate to high accuracy for non-destructive assessment of cassava below-ground traits. Initial studies showed strong correlations between GPR signals and measured root biomass, with correlation coefficients around  $r \approx 0.79$  and  $R^2$  values up to ~0.63–0.77 depending on genotype [35]. Subsequent analyses using improved signal processing and predictive models achieved cross-validated prediction accuracies of  $r \approx 0.64$ –0.72 for root biomass and early bulking indicators [35,70]. These results demonstrate that GPR

can reliably capture root volume, architecture, and early bulking dynamics, supporting its use as a high-throughput phenotyping tool in cassava breeding programs.

Yield-related traits remain among the hardest to dissect genetically in cassava due to their strong polygenic control and susceptibility to environmental variation. Studies such as Abah et al. (2024) [71] highlight that QTLs for early bulking, bulking rate, and root size display strong genotype  $\times$  environment interactions. QTLs identified in one environment often have weak or inconsistent effects elsewhere. Traits such as harvest index, stem number, canopy development, and root number are governed by many small-effect loci distributed across the genome [11,16].

Because bulking and yield traits are highly polygenic, they are ideal targets for genomic selection, which captures genome-wide additive variation rather than relying on individual marker effects.

## THE CASSAVA GENOME AND FUNCTIONAL GENOMICS

Advances in cassava genomics over the past decade have radically improved our understanding of the crop's biology and opened new frontiers for genetically informed breeding. The availability of a chromosome-level reference genome, coupled with functional genomic tools such as transcriptomics, epigenomics, and regulatory network analyses, allows researchers to dissect complex traits, including early storage root formation, disease resistance, and starch quality, at molecular resolution. These resources have significantly strengthened the integration of genomics into cassava breeding programs.

### The Cassava Genome

The release and continuous improvement of the cassava reference genome have been transformative for cassava biology. The current assembly (v8.1) provides one of the most complete and accurate representations of the cassava genome to date, offering  $\sim$ 750 Mb genome assembly, capturing the majority of euchromatic and heterochromatic regions; approximately 33,000 predicted protein-coding genes, reflecting cassava's high allelic diversity and genomic complexity; chromosome-level scaffolding, enabling precise gene localization, structural variant identification, and comparative mapping across *Manihot* species [14,32]. This high-quality assembly supports diverse downstream applications, including GWAS, QTL mapping, genomic selection, and evolutionary studies.

### Key Gene Networks Uncovered Through Genomic Analyses

The reference genome has significantly improved gene discovery for multiple biochemical and physiological pathways central to cassava improvement:

In starch biosynthesis pathway, genes underpinning starch production and functionality, including *granule-bound starch synthase I (GBSSI)*, *starch branching enzymes (SBE1/2)*, and *soluble starch synthase II (SSII)*, have been fully annotated. These genes explain variation in amylose content, pasting viscosity, gelatinization temperature, and suitability for industrial applications [72,73].

In cyanogenic glucoside pathway, the biosynthesis of linamarin and lotaustralin is governed by cytochrome P450 genes *CYP79D1*, *CYP79D2*, and downstream enzymes such as hydroxynitrile lyase (HNL). These genes have become major targets for engineering low-cyanide cassava and improving food safety [74].

Study on carotenoid accumulation, showed that biofortification efforts have benefited greatly from genomic identification of genes regulating  $\beta$ -carotene levels, including *phytoene synthase (PSY)*, *orange (OR)*, and  *$\beta$ -carotene hydroxylase*. These loci underpin the development of yellow-fleshed cassava with enhanced vitamin A content [75,76].

Under the root initiation and bulking regulatory networks, genomic analyses have revealed transcription factors, hormone signaling components (auxin, cytokinin, ABA), and cell-wall modifying enzymes that influence root initiation, secondary growth, and early bulking. Key regulators include MCM1, AGAMOUS, DEFICIENS, SRF (MADS)-box genes, NAM (No Apical Meristem), ATAF1/2, CUC2 (NAC)-domain proteins, expansins, and sugar transporters, elements crucial for early storage root formation. Together, these genomic discoveries provide unprecedented opportunities to manipulate trait pathways through MAS, genomic selection, and targeted gene editing [11,16,19].

## Transcriptomics

Transcriptomics, especially RNA sequencing (RNA-seq), has become one of the most powerful tools in cassava functional genomics. By profiling gene expression under different developmental stages, environments, and stress conditions, researchers can elucidate the molecular mechanisms regulating complex traits.

### Insights from Cassava Transcriptome Studies

**On Early Storage Root Initiation Pathways;** RNA-seq analyses comparing fibrous roots, initiating storage roots, and bulking roots reveal shifts in expression of genes associated with: cambial activation; starch biosynthesis; sugar transport [Sugars Will Eventually be Exported Transporters (SWEET) and Sucrose Transporters (SUT)]; phytohormone biosynthesis and signaling. These expression signatures highlight candidate genes that may control the timing and intensity of early bulking, a critical breeding target [41].

**Drought Response Regulatory Modules:** Cassava exhibits remarkable resilience to episodic drought. Transcriptomic profiling under water stress conditions has identified: ABA-responsive transcription factors; late-embryogenesis abundant (LEA) proteins; ROS-scavenging enzymes; aquaporins and osmotic adjustment proteins. These modules reveal how cassava maintains carbon assimilation and root growth under fluctuating moisture regimes [77,78].

**CMD Resistance Transcriptional Networks:** Comparative transcriptomics between CMD-resistant and susceptible genotypes show differential expression of: NBS-LRR resistance genes; RNA silencing components; viral replication inhibitors; cell wall-associated defense genes. Insights from these studies complement genetic mapping of CMD2/CMD3 and help clarify the molecular basis of durable resistance [79–81].

**Postharvest Physiological Deterioration (PPD) and Cooking Quality:** PPD, a major challenge affecting fresh root shelf life, is associated with rapid oxidative bursts and activation of stress-related genes. Transcriptomic studies have identified: peroxidases; phenolic pathway enzymes; programmed cell death regulators. Cooking quality studies show differential expression of genes controlling cell wall softening, starch gelatinization, lignification, and water binding capacity, all critical in traits such as fufu texture, gari expansion, and boiling time [82,83].

### ***Transcriptomics as a Breeding Accelerator***

The integration of transcriptomic data with genomic markers, metabolomics, and phenomics enables: identification of predictive expression markers; mapping of regulatory networks; development of gene-based selection indices; improved genomic prediction accuracy; discovery of candidate genes for Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-based editing.

Thus, transcriptomics now plays a central role in modern cassava improvement strategies, bridging the gap between genotype and phenotype [84].

## **GENOME-WIDE ASSOCIATION STUDIES IN CASSAVA**

Genome-wide association studies have become one of the most useful tools in cassava genetics, enabling the dissection of complex, quantitatively inherited traits at high resolution. Cassava's reproductive biology, namely, its obligate outcrossing system and high heterozygosity, results in rapid linkage disequilibrium decay across the genome. This unique genomic characteristic allows GWAS to pinpoint loci with greater precision than in many other clonally propagated crops [50]. The widespread adoption of high-density SNP genotyping platforms such as GBS, DArTseq, and GEM SNP chips has further expanded the power of GWAS across diverse breeding populations.

GWAS has now been applied to a wide range of agronomic, nutritional, physiological, and disease-resistance traits, uncovering genetic variants of direct relevance to breeding programs at CIAT, IITA, NRCRI, EMBRAPA, and other global partners [17,23,71].

### Why GWAS Works Well in Cassava

Several genomic and biological features make cassava particularly suited for GWAS:

As a result of low to moderate LD or rapid LD decay; LD in cassava declines rapidly over short genomic distances due to frequent outcrossing. This allows researchers to identify loci with fine-scale resolution, narrowing candidate gene lists dramatically [49].

As a result of high allelic diversity; extensive diversity in landraces and breeding populations increases the probability of capturing functional alleles associated with traits of interest [85,86].

Availability of large, structured breeding populations; Multi-family, multi-environment datasets developed by CIAT, IITA, and NextGen Cassava provide robust phenotypic and genotypic resources for association mapping [87].

As a result of chromosome-level reference genome; The V8.1 assembly allows precise anchoring of significant SNPs and facilitates candidate gene discovery.

Combined, these features support high-confidence mapping of traits governed by many small-effect loci, traits traditionally considered “intractable” under classical breeding.

### Key GWAS Discoveries in Cassava

#### ***β*-Carotene Accumulation and Nutritional Biofortification**

GWAS studies have successfully identified major loci associated with carotenoid accumulation, crucial for developing biofortified cassava to combat vitamin A deficiency. Ikeogu et al. (2009) [88] identified SNPs linked to *PSY2*, *ORANGE (OR)*, and downstream carotenoid pathway genes. These discoveries enabled genomic selection models that predict carotenoid content with high accuracy and supported marker-assisted selection for yellow-fleshed varieties. These markers have already been integrated into biofortification pipelines at IITA and HarvestPlus.

#### **Disease Resistance (CMD and CBSD)**

Genome-wide association studies have significantly advanced the molecular understanding of cassava’s two most devastating viral diseases: Cassava Mosaic Disease and Cassava Brown Streak Disease. For CMD, GWAS has consistently identified strong associations on Chromosome 12, corresponding to the well-characterized CMD2 locus, as well as additional quantitative resistance regions on Chromosome 16 [17]. These loci have been shown to contribute to durable, broad-spectrum

resistance, complementing earlier QTL mapping studies and providing molecular targets for marker-assisted selection. Analyses of these regions reveal enrichment of functional transcripts related to key defense mechanisms, including Nucleotide-Binding Site-Leucine-Rich Repeat (NBS-LRR) resistance genes, components of the RNA-silencing machinery, and other antiviral effectors. Such genes form part of the cassava innate immune system, enabling recognition of viral pathogens and activation of downstream defense responses. In contrast, CBSD resistance exhibits a more polygenic architecture and is highly influenced by environmental conditions, making it more complex to dissect genetically. Nevertheless, GWAS analyses have successfully identified consistent SNP markers associated with CBSD resistance, often located near genes implicated in cell-wall signaling, antiviral responses, and root-specific defense mechanisms [89]. These findings highlight the role of multiple, small-effect loci in conferring partial resistance and underscore the importance of integrating genomic information across diverse environments for breeding CBSD-resistant varieties.

### ***Starch Quality, Dry Matter Content, and Physicochemical Traits***

GWAS has identified SNPs associated with starch composition, including amylose:amylopectin ratios, starch branching enzymes, granule size, dry matter content, and hydration/pasting behavior [16,90]. These loci are directly relevant to industrial processing (starch, ethanol), food-texture traits (fufu moldability, gari expansion), and improved shelf stability. The findings enable breeders to select for complex quality traits, integrating functional genomics with phenotypic data to enhance both product quality and processing efficiency.

### ***Early Storage Root Bulking***

Emerging GWAS analyses have begun revealing early bulking markers associated with: sugar transport (SWEET gene family); cambial activity regulators; hormone signaling pathways; carbohydrate partitioning networks.

Although effect sizes are typically small and environmentally variable, the identification of these markers provides useful biological insight and supports development of genomic prediction models tailored for early bulking [71].

### ***Multi-Environment GWAS: A Major Advancement***

Cassava traits often display strong genotype  $\times$  environment (G  $\times$  E) interactions, especially disease scores, root quality, and bulking traits. Single-environment GWAS analyses frequently fail to detect stable, robust loci.

### ***Benefits of Multi-Environment GWAS Approaches***

Combining multi-environment datasets has brought several key benefits for cassava breeding:

It helps to increased power and reliability; by combining phenotypic and genotypic data across diverse locations, researchers can detect loci that remain stable under variable disease pressure, climate, and soil conditions. This approach reduces environmental noise and improves the confidence in identified marker–trait associations. Stable loci across environments provide reliable targets for selection [91].

It helps in the identification of environment-specific alleles; multi-environment trials reveal alleles that are expressed or show significant effects only under certain stresses or agro-ecological conditions, such as drought, nutrient limitation, or high disease prevalence. These environment-specific alleles capture adaptive responses and inform breeding for stress resilience [17].

It provides improved integration with genomic prediction; loci that demonstrate stability across environments contribute consistent signals to genomic prediction models, increasing the accuracy of genomic estimated breeding values (GEBVs) for traits of interest. Multi-environment GWAS and mixed-model approaches that account for genotype  $\times$  environment interactions enhance predictive performance compared with single-environment analyses. [22,92].

### **GWAS as a Breeding Accelerator**

The integration of GWAS into breeding pipelines has produced several practical outcomes: enables discovery of candidate genes for functional validation; supports MAS for major-effect loci (CMD2, carotenoid genes); feeds trait-associated SNPs into genomic selection models; refines breeding population design and parental selection; accelerates genetic gain for polygenic traits [17,22,70].

With increasing availability of multi-omic data, transcriptomics, metabolomics, and phenomics, future GWAS approaches may adopt multi-omic GWAS and causal modeling to uncover deeper regulatory layers.

## **GENOMIC SELECTION (GS) IN CASSAVA BREEDING**

Genomic selection (GS) has emerged as one of the most transformative innovations in cassava improvement, enabling breeders to predict the genetic potential of individuals using genome-wide markers before phenotypic evaluation is completed. This approach is particularly impactful for cassava, where long growth cycles, clonal propagation, and strong genotype  $\times$  environment (G  $\times$  E) interactions slow down gains from traditional breeding. By capturing effects from thousands of small-effect markers simultaneously, GS accelerates the identification of superior clones and supports rapid-cycle breeding pipelines.

GS was first implemented in cassava through the collaborative efforts of the NextGen Cassava Breeding Project, integrating genomic datasets from CIAT, IITA, NRCRI, and national programs across Africa and Latin America. Today, GS is fully embedded in elite breeding pipelines and has demonstrably increased genetic gains for yield, resilience, and quality traits [17].

### **Applications of GS in Cassava**

#### ***Root Yield and Productivity Traits***

Yield is one of the most complex traits in cassava, controlled by numerous loci of small effect and heavily influenced by environmental variation. Traditional QTL-based approaches explain only a small portion of yield variance. GS overcomes this limitation by using all genome-wide SNPs to predict breeding values.

Wolfe et al. (2017) [11] demonstrated that GS achieves moderate-to-high prediction accuracy for root yield, harvest index, and related productivity traits.

GS enables selection among seedlings or early-stage clones without waiting for full clonal evaluations, dramatically speeding up genetic gains.

Yield-related GS models have been particularly effective when trained on multi-year, multi-location datasets, allowing better modeling of  $G \times E$  patterns.

#### ***Cassava Mosaic Disease (CMD) Resistance***

Although the major loci CMD2 and CMD3 confer strong resistance to cassava mosaic disease (CMD), the observed phenotypic expression of CMD resistance still exhibits quantitative variation among progenies. This variation arises from the contributions of multiple minor-effect loci, background genetic variation, and genotype  $\times$  environment interactions. Genomic selection (GS) has become a key tool in breeding programs to manage this complexity. By using genome-wide marker data, GS enables the prediction of CMD resistance in progenies even before they experience significant disease pressure, allowing early and accurate selection of promising clones. Additionally, GS facilitates the combination of major-effect CMD loci with minor background resistance alleles, supporting the development of durable and broad-spectrum resistance that is less vulnerable to pathogen evolution. Another advantage of GS is that it helps resolve confounding effects between CMD resistance and growth-related traits during clonal selection, allowing breeders to select for resistance without inadvertently selecting against yield or other agronomic traits. This integration of genome-wide information with breeding strategies has been demonstrated to improve selection efficiency and accelerate genetic gain for CMD resistance in cassava breeding programs [93].

### ***Root Quality Traits (Dry Matter, Starch, Carotenoids)***

Quality traits are essential for consumer preference, processing properties, and nutrition, show high heritability and strong genetic signals. GS has proven effective for: dry matter content (DM) as a key trait for gari, fufu, and industrial starch; starch viscosity and physicochemical properties; carotenoid content, enabling rapid advancement of biofortified yellow cassava. Combining genomic and metabolomic data further improves prediction accuracy for complex starch and quality traits [94].

### ***Multi-Environment Genomic Predictions***

Cassava is cultivated across diverse agro-ecologies with varying disease pressure, rainfall patterns, soil fertility, and climate regimes. To address these challenges: Multi-environment GS models incorporate  $G \times E$  interactions directly into predictions; Reaction-norm models, genomic genotype-by-environment models, and Bayesian hierarchical approaches significantly improve prediction accuracy; These models have been particularly valuable for yield, CBSD responses, and drought-related traits. This approach strengthens varietal stability and supports more targeted deployment recommendations [95].

### ***Benefits of Genomic Selection***

GS offers several strategic advantages that align exceptionally well with cassava's biology and breeding system.

#### ***Shortened Breeding Cycles (Reduction by 2–4 Years)***

Cassava's long growth cycle traditionally limits breeding progress. GS: allows early selection at seedling or nursery stages; reduces the number of clonal evaluation cycles; enables rapid-cycle recurrent selection. Breeding programs that once completed a cycle in 8–10 years can now complete cycles in 4–6 years.

#### ***Increased Selection Intensity***

Because GS allows selection from large seedling populations without the need for extensive early phenotyping: breeders can screen thousands of genotypes cheaply using DNA alone; only the genetically superior clones advance to field testing; selection intensity increases substantially, improving genetic gain per cycle. This is especially important for elite  $\times$  elite crosses targeting yield, CMD resistance, and root quality.

#### ***Improved Accuracy for Polygenic Traits***

Traits such as yield, bulking rate, dry matter content, and stress resilience are governed by many loci with small individual effects. GS excels in these contexts because: it captures genome-wide additive and non-additive variance; it avoids the limitations of QTL-based MAS; it

improves prediction accuracy even when individual markers explain small portions of variance. Thus, GS is a natural fit for cassava's most challenging traits.

### ***Seamless Integration with Clonal Propagation***

Cassava's vegetative propagation system supports GS in several ways: once superior clones are identified via genomic prediction, they can be propagated and tested quickly; genomic predictions made early in the pipeline directly translate into clonal performance; multi-stage clonal trials can be reorganized to focus resources on genetically superior candidates. This integration enhances both efficiency and resource use.

### ***Supports Modern Breeding Objectives***

GS also complements other cutting-edge innovations like high-throughput phenotyping (HTP) for root architecture and canopy traits; multi-omic prediction models that combine SNPs, transcripts, and metabolites; gene editing by identifying high-value genetic targets; accelerated parental selection enabling more frequent crossing cycles. Through these synergies, GS strengthens the entire cassava breeding ecosystem.

## **MULTI-OMICS INTEGRATION IN CASSAVA IMPROVEMENT**

Recent advances in high-throughput “omics” technologies have opened new possibilities for understanding cassava biology at molecular, biochemical, and systems levels. Because cassava traits such as storage root bulking, disease resistance, postharvest quality, and nutritional attributes are multi-layered and polygenic, single-omic approaches often provide only partial insight. Multi-omics integration, combining metabolomics, transcriptomics, proteomics, genomics, and phenomics, offers a holistic framework to uncover regulatory mechanisms and accelerate precision breeding.

These tools enable breeders and biologists to link genotype to phenotype more accurately, identify biomarkers, model regulatory networks, and predict complex trait performance with greater confidence. Multi-omics approaches are now central to modern cassava improvement programs, particularly through initiatives such as NextGen Cassava, HarvestPlus, CGIAR Excellence in Breeding, and national programs at NRCRI, IITA, and CIAT [96].

### **Metabolomics**

Metabolomics has emerged as a strategic molecular phenotyping platform in cassava research because metabolites represent the end-products of biological regulation and therefore define the direct biochemical phenotype of a genotype. Unlike DNA markers or transcripts, which indicate potential or regulatory activity, metabolites reflect the

realized physiological and biochemical state, providing the most proximal measurable layer linking gene function to expressed agronomic, nutritional, and processing traits [97]. Cassava storage roots are biochemically rich sink organs characterized by large carbon flux, starch accumulation, cyanogenic compounds, soluble sugars, and wide metabolic plasticity, making metabolomics particularly well-suited for interrogating complex quality traits, toxicity reduction, stress physiology, and end-use properties relevant to consumer and industrial priorities [14,98].

Transcriptome-to-metabolome associations have shown that metabolic traits can vary even among genotypes that share similar QTL profiles, due to rapid post-transcriptional regulation and environmental responsiveness [99]. This reinforces the importance of metabolomics for breeding decisions in cassava improvement pipelines, most especially for traits where biochemical composition defines product identity more strongly than agronomic appearance alone [100]. Cassava breeding programs now increasingly rely on metabolomic readouts to capture traits such as starch behavior, sugar partitioning, cyanide potential, volatile compound diversity, root aging physiology, and stress-induced metabolic resilience, all of which are critical to end-user acceptability and variety adoption [101].

One of the most impactful areas of metabolomics application in cassava is the characterization of storage-root starch structure and its physicochemical behavior under processing. By profiling biosynthetic intermediates associated with amylose and amylopectin assembly, metabolomics enables differentiation of starch granule branching density, chain-length distribution, crystalline lamella arrangement, peak viscosity, pasting temperature, and gelatinization dynamics. These parameters directly influence sensory texture and final product behavior in major cassava-based foods, including fufu, gari and industrial starch forms [101,102]. Metabolomic profiles showing elevated root-localized carbon intermediates of the ADP-Glucose pathway have consistently correlated with increased storage-root bulking efficiency and improved starch cohesive behavior during dough formation, key parameters that determine fufu moldability, elasticity, and product texture perception [103]. These biochemical measurements support breeders in selecting genotypes that produce starch with optimal swelling, dispersibility, retrogradation patterns, and paste stability required for both fermented cassava foods and starch-industry utility [103].

Metabolomics equally supports quantitative resolution of sugar profiles, offering insight into carbon partitioning efficiency and sink strength in storage-root initiation. Comparative carbohydrate pool analyses (glucose, fructose, sucrose, and maltose levels) have allowed characterization of carbon allocation dynamics between source tissues (leaves) and sink organs (roots). Genotypes with high sucrose abundance in storage-root initials exhibit higher sink strength, increased phloem

unloading efficiency, and accelerated root conversion from fibrous to bulking state, thereby promoting earlier bulking initiation and increased fresh-root biomass yield [104]. Stable sugar assays using Gas Chromatography–Mass Spectrometry (GC-MS) have been widely validated as biochemical predictors for selecting cassava lines that demonstrate optimal carbon flow into roots and yield-stability under fluctuating climatic or agronomic environments, offering critical value in juvenile-phase selection during breeding cycles [100].

Efforts to reduce cassava toxicity have also been advanced by metabolomics through direct measurement of cyanogenic glucosides, primarily linamarin and lotaustralin. These aliphatic nitrile-linked secondary metabolites represent the major cyanide sources in cassava, and their quantification by metabolomic platforms allows for high-precision estimation of total cyanide potential (TCP) across genotypes, environments, and processing states. This precision is essential when breeding for varieties suitable for direct human consumption or fast processing, where cyanide thresholds determine safety compliance [98,101]. In West African breeding scenarios, metabolomic cyanide mapping has significantly improved the ability of breeding programs to prioritize low-linamarin genotypes without sacrificing root size or dry-matter content, a balance that is otherwise difficult to achieve using agronomic selection alone [14,105]. These safety-guided metabolite screens are now viewed as essential complementary analyses alongside field-based dry-matter or fresh-root yield data.

Metabolomics has also progressed the molecular interpretation of consumer-driven quality traits. Storage-root metabolite profiling has supported characterization of volatile organic compounds (VOCs), short- and long-chain fatty acids, organic acids (lactic, malic, citric, succinic), aldehydes, esters, ketones, alcohols, phenolics and flavor-conferring secondary metabolites that collectively define taste, aroma, after-fermentation acidity, dough collapse resistance and textural perception. These biochemical traits influence fufu aroma intensity, elasticity, dough stretchability, moldability, stickiness, taste complexity and user-perceived cooking quality [101]. Sensory metabolomics studies showed that cassava genotype-specific VOC clusters can distinguish consumer-preferred organoleptic profiles for fermented cassava food products, supporting biochemical trait selection for variety-adoption breeding cycles [100]. Moreover, metabolomics-sensory correlation models allow screening for secondary metabolites (fermentation-related volatiles, polyols, phenols) that both enhance eating quality and minimize off-aromas or textural collapse during dough shaping, helping breeding programs better predict gastronomic acceptability through biochemical fingerprinting rather than field appearance alone [99].

Because metabolites respond dynamically to both genotype and the environment, metabolomics equally supports modeling of genotype  $\times$  environment  $\times$  metabolite (G  $\times$  E  $\times$  M) interactions. These interactions enhance understanding of environment-sensitive processes including early bulking response to rainfall or dry seasons, metabolite fluctuation under drought or virus pressure, browning sensitivity after harvest, and metabolite-induced variability during fermentation-or-processing phases. The ability to characterize these G  $\times$  E  $\times$  M interactions allows breeding programs to identify metabolic stability markers for robust varietal screening under climate uncertainty [41,104]. Thus, beyond quality trait characterization, metabolomics supports cassava breeding by monitoring biochemical consistency, root aging physiology, food-metabolite compliance, stress memory pathways, and metabolic vigor that translate into reliable enterprise performance for women-led processing units dependent on cassava product quality in real-world settings [106,107].

Metabolomics provides cassava breeding with the ability to fingerprint biochemical identity of storage-root starch or sugar architecture, measure cyanide potential to ensure product compliance, characterize VOC and flavor-conferring molecule clusters that determine fufu or gari quality, and model G  $\times$  E  $\times$  M systems for climate-responsive variety stability. These contributions push cassava breeding beyond indirect markers toward biochemical phenotype-first varietal improvement models that enhance both consumer-preferred qualities and industrial starch adoption [99].

### Transcriptomics

Transcriptomics, largely enabled by RNA sequencing technology (RNA-seq), offers genome-wide resolution of transcriptional activity, allowing researchers to quantify and compare gene expression dynamics across tissues, developmental stages, environmental conditions, and contrasting genotypes. In cassava, transcriptomic resources have played a foundational role in decoding regulatory circuits underlying storage-root formation, disease resistance, abiotic stress responses, and nutritional improvement. These datasets bridge molecular biology and breeding by revealing not only differentially expressed genes but also gene co-expression modules, pathway activation hierarchies, and regulatory hubs that coordinate complex trait expression [99].

One of the earliest applications of transcriptomics in cassava bulking research focused on distinguishing fibrous (non-storage) roots from storage-root initials. RNA-seq comparisons demonstrated that early bulking is marked by upregulation of gene families involved in cambial and meristematic reactivation, including members of the MADS-box, NAC, and WOX transcription factor families, which promote vascular cell identity and secondary root thickening. These studies further revealed heightened transcription of sucrose transporters (including SUT and SWEET gene clades), starch assembly enzymes, and hormone-responsive

regulators integrated in auxin, abscisic acid (ABA), and cytokinin signaling pathways, collectively defining the molecular shift from exploratory to storage sink behavior [41]. The identification of these bulking-associated expression signatures has provided a template for modeling source-to-sink transitions and for creating transcriptional biomarkers that inform early selection in breeding populations [41].

Cassava mosaic disease resistance has also been clarified through differential transcriptome analyses between resistant genotypes carrying CMD2 or CMD3 loci and susceptible lines. These studies showed strong transcriptional enrichment of activation-induced immune receptors, particularly NBS-LRR (nucleotide-binding site leucine-rich repeat) and LRR-RLK (receptor-like kinase) sensors, which activate patterned antiviral signaling. Transcriptomics also uncovered high expression of RNA interference (RNAi) pathway genes, including DCL (Dicer-like), RDR (RNA-dependent RNA polymerase), andAGO (Argonaute) families that regulate antiviral small-RNA biogenesis. In addition, defense-primed genes contributing to pectin methylesterification, callose deposition, and cell-wall fortification were observed to support resistance containment and durability [108,109]. These mechanistic findings have complemented trait loci detected through QTL or GWAS by explaining how resistance alleles shape downstream immunity cascades for more durable viral defense [104].

Under drought and prolonged water-limitation conditions, cassava mounts a layered transcriptional response dominated by ABA-responsive regulons. Transcriptome studies identified dehydration-induced activation of hormonal receptors [Pyrabactin Resistance/ PYR1-Like/ Regulatory Component of ABA Receptor (PYR/PYL/RCAR)], transcriptional regulators [ABA-Responsive Element Binding protein/ ABRE-Binding Factor (AREB/ABF)], osmolyte and osmoprotectant biosynthesis pathways (including raffinose, trehalose, and late-embryogenesis abundant proteins), and antioxidant enzyme gene clusters [Superoxide Dismutase (SOD), Catalase (CAT), Ascorbate Peroxidase (APX), Glutathione Peroxidase (GPX)] that protect cells from induced oxidative burst. WRKY, Myeloblastosis (MYB), Basic Leucine Zipper (bZIP), and Basic Helix-Loop-Helix (bHLH) transcription factor families were consistently reported as central regulators of stomatal modulation, cellular homeostasis, sugar reallocation, and gene network reinforcement under drought stress. These integrated pathways explain the molecular basis of cassava's unique ability to maintain productivity under low water environments [104,110].

Cassava biofortification efforts, particularly breeding for enhanced  $\beta$ -carotene content, have equally benefited from transcriptomics. Deep expression profiling of carotenoid pathways revealed structural gene upregulation, including phytoene synthase (PSY), orange (OR) chaperones, carotenoid cyclases [Lycopene  $\beta$ -Cyclase/ Lycopene  $\epsilon$ -Cyclase (LCYB/LCYE)], and plastid-localized transport regulators that influence

carotenoid sequestration and stability. Weighted gene co-expression network analysis (WGCNA) has identified metabolic modules linking PSY-OR co-activation with downstream carotenoid assembly networks, providing clear path for nutritional trait prediction and marker development for biofortified breeding pipelines [111]. The ability to decipher such co-regulated networks enables breeders to shift away from single-gene models toward coordinated pathway-level improvement, which is especially important for heterozygous crops such as cassava [99].

Collectively, transcriptomics has empowered cassava breeding by enabling: identification of developmental switches in early storage-root bulking; clarification of defense regulons underlying CMD resistance; resolution of hormone-driven drought stress signaling networks; and mapping of pathway-co-expression modules that improve nutritional biofortification. These advances demonstrate that cassava improvement increasingly depends on multi-layered gene network interpretation rather than isolated genes, providing the biological context necessary to design superior genomic selection models for root yield, disease durability, industrial starch profile, and climatic resilience [99,104].

### Proteomics

Proteomics provides insight into functional proteins, the molecular machines executing biochemical pathways. Because proteins undergo post-translational modifications and are not always predictable from transcript levels, proteomics fills critical gaps left by other omics layers. Proteomics has made significant contributions to understanding cassava biology, especially with respect to root development, starch metabolism, hormonal regulation, and stress responses. For instance, comparative proteomic studies during cassava storage-root development have identified many enzymes involved in carbohydrate metabolism and starch accumulation, including active isoforms of starch-branching enzymes and other starch-synthesis related proteins, underscoring the biochemical control of starch architecture in developing roots [31]. Beyond starch metabolism, proteome analyses of cassava roots at different developmental stages have revealed shifts in proteins linked to energy production, cell structure, and stress defense, illustrating the complexity of metabolic reprogramming during storage-root formation [112].

Moreover, proteomic investigations under abiotic stress conditions, for example drought, have documented alterations in the abundance of stress- and defense-related proteins, including antioxidant enzymes and proteins involved in ROS detoxification. These findings highlight potential molecular mechanisms by which cassava adjusts its physiology under water deficit, offering candidate targets for breeding drought-resilient varieties [110]. Proteomics thus provides a functional, protein-level view of the cellular machinery underlying root growth,

starch deposition, and stress responses in cassava, insights that go beyond gene expression studies and help bridge genotype to phenotype for traits of agronomic importance such as yield, starch quality, and stress resilience.

### **Integrated Platforms: Multi-Omics Fusion for Precision Breeding**

The power of multi-omics lies not in individual datasets but in their integration. When genomics, metabolomics, transcriptomics, proteomics, and phenomics are fused within systems biology frameworks, they enable unprecedented clarity into the genotype–phenotype continuum.

#### ***Benefits of Multi-Omics Integration***

The integration of multi-omics datasets offers several significant benefits for cassava breeding and the understanding of complex traits. One major advantage is biomarker discovery, as integrative analyses can identify metabolites, genes, or protein signatures that are predictive of traits such as early bulking, post-harvest physiological deterioration (PPD) resistance, cassava mosaic disease (CMD) immunity, or cooking quality. These biomarkers can accelerate breeding decisions and enhance the accuracy of selection by providing measurable indicators of desired phenotypes.

Multi-omics integration also facilitates the identification of regulatory hubs within biological networks. Network modeling can reveal master regulators, including transcription factors, signaling molecules, and hormonal nodes that control multi-gene pathways. These hubs represent prime candidates for gene editing, offering targeted opportunities to influence complex trait expression.

Moreover, integrating omics datasets can provide mechanistic insights into complex traits by uncovering causal pathways that link gene expression, metabolic flux, and phenotypic outcomes. This understanding is particularly critical for polygenic traits, where the interplay of multiple genes and pathways determines the final phenotype.

Enhanced genotype–phenotype associations are another key benefit. Approaches such as multi-omic genome-wide association studies (GWAS) and multi-kernel genomic prediction models combine multiple layers of biological information, improving the ability to detect trait-associated loci and increasing prediction accuracy.

Multi-omics integration opens new pathways for climate-resilient breeding. Traits such as drought tolerance, heat resilience, and disease resistance can be modeled as interconnected, multi-layered biological systems rather than as isolated features. This holistic view enables the development of more robust cassava varieties that are better adapted to environmental stresses, supporting sustainable crop production in the face of climate change.

## FUTURE PROSPECTS FOR GENOMICS-ASSISTED CASSAVA BREEDING

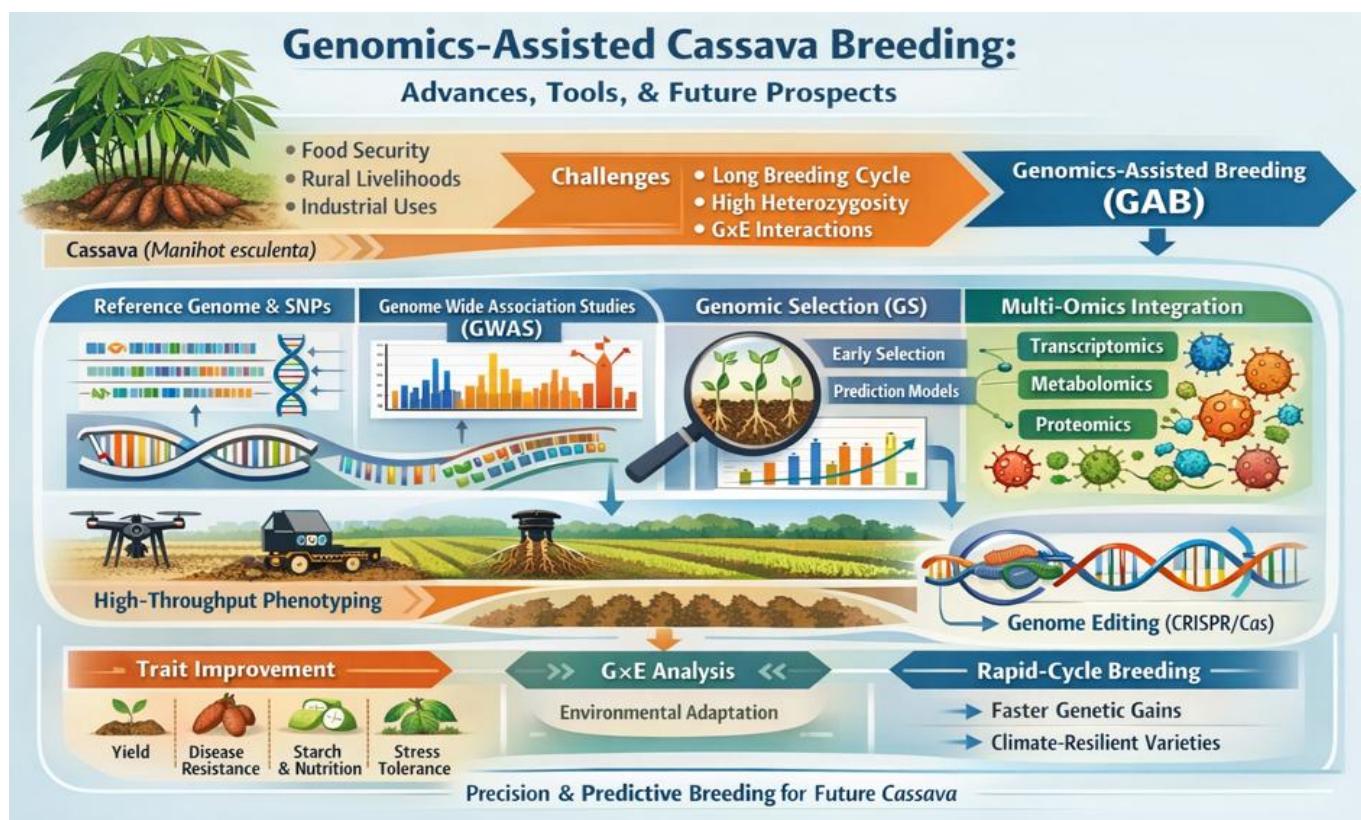
The rapid development of genomic tools and biotechnologies promises to transform cassava breeding, enhancing the crop's potential to contribute to food security, nutrition, climate resilience, and socio-economic development. One of the key drivers of future cassava improvement is the application of genomic selection (GS) and marker-assisted selection (MAS), facilitated by the increasing availability of high-density molecular markers such as SNPs and SSRs, as well as high-throughput genotyping and comprehensive genome data [113,114]. Genomic selection enables breeders to predict breeding values based on genomic information, reducing the need for lengthy and costly field trials, a particularly important advantage given cassava's long growth cycle and highly heterozygous genome [113]. Large-scale genome resequencing efforts, which have analyzed hundreds of cassava accessions and identified thousands of high-quality SNPs and candidate loci for critical agronomic traits, further enable mapping and selection for complex traits such as yield stability, root quality, stress tolerance, and heterozygosity management [114].

Advances in gene editing and molecular biotechnology also offer unprecedented opportunities for targeted trait improvement in cassava. The advent of precise genome editing tools, most notably CRISPR/Cas9, allows the development of cassava varieties with enhanced disease resistance, improved nutritional profiles, superior starch quality, reduced cyanogenic glucoside content, and delayed post-harvest deterioration, traits that are difficult or slow to achieve through conventional breeding [114]. Gene editing can enable the removal or modification of susceptibility genes or metabolic pathways to reduce cyanide content or improve starch composition [114]. Complementing these techniques, "omics"-based functional genomics, including transcriptomics, proteomics, and metabolomics, provides detailed insights into the molecular basis of stress tolerance, root quality, starch biosynthesis, and other complex traits, thereby supporting more precise and informed breeding decisions [113,114]. These integrated genomic and functional approaches are likely to be pivotal in developing cassava varieties that meet both agronomic requirements and consumer demands.

In addition, new breeding strategies are expected to shorten breeding cycles and increase genetic gain. Inbred-parent-based hybrid breeding strategies have been proposed to overcome challenges related to heterozygosity and long generation times [113]. When combined with genomic selection and advanced genotyping, these strategies could significantly reduce breeding cycles, improve trait introgression, eliminate deleterious alleles, and exploit heterosis. Such approaches promise to accelerate the development of cassava varieties that are resilient to climate variability, resistant to emerging pests and diseases, and better aligned with evolving market demands [114].

The effectiveness of genomics-assisted breeding also depends on robust data management and high-throughput phenotyping. Digital platforms such as CassavaBase allow breeding programs to efficiently manage large datasets, including phenotyping, genotyping, and crossing information, thereby improving quality control and accelerating decision-making [115]. High-throughput phenotyping methods, including near-infrared spectroscopy, drone- or imaging-based field assessment, and automated root quality evaluation, are likely to become increasingly widespread [116]. These technologies help overcome bottlenecks associated with slow, labor-intensive, and error-prone manual phenotyping, particularly for complex traits such as yield stability, drought tolerance, root quality, and post-harvest resilience.

Collectively, these advances suggest a future in which genomics-assisted cassava breeding can deliver high-yielding, nutrient-rich, and climate-resilient varieties more rapidly and efficiently than ever before. By integrating genomic tools, molecular biotechnology, advanced breeding strategies, and digital phenotyping platforms, cassava breeding programs are positioned to transform both the crop and the livelihoods of the millions who depend on it, particularly in regions vulnerable to food insecurity and climate stress [113,114] as shown in Figure 1.



**Figure 1.** The future of quality cassava production through precision and predictive breeding.

## CONCLUSIONS

Genomic selection (GS) and multi-omics integration are transforming cassava breeding by providing tools to accelerate genetic gain, improve accuracy in complex trait prediction, and optimize resource allocation. GS enables earlier selection at seedling or nursery stages, shortens breeding cycles by 2–4 years, and increases selection intensity, particularly for polygenic traits such as yield, dry matter content, and stress resilience. Its seamless integration with clonal propagation ensures that genetic predictions translate directly into field performance, allowing breeders to focus resources on the most promising genotypes.

The integration of metabolomics, transcriptomics, proteomics, and other omics layers provides a comprehensive view of cassava biology, linking genotype to phenotype with unprecedented precision. Metabolomics enables the profiling of storage-root biochemical traits, cyanogenic glucoside content, starch architecture, sugar partitioning, and volatile organic compounds that determine consumer-preferred qualities and processing suitability. Transcriptomics reveals regulatory networks controlling storage-root initiation, drought tolerance, CMD resistance, and biofortification traits, while proteomics uncovers functional proteins that execute biochemical pathways and stress responses.

Multi-omics integration enhances the discovery of biomarkers, regulatory hubs, and causal pathways, thereby improving selection accuracy and enabling precision breeding for complex traits. These approaches also support the development of climate-resilient cassava varieties, capable of maintaining yield, quality, and stress tolerance under fluctuating environmental conditions. Advances in gene editing, high-throughput phenotyping, and digital breeding platforms further complement these innovations, offering the potential to accelerate breeding cycles, reduce labor-intensive field evaluation, and optimize trait introgression.

The combination of GS, multi-omics, and modern breeding strategies positions cassava improvement programs to meet both agronomic and industrial demands. These approaches facilitate the development of high-yielding, nutritious, safe, and climate-resilient cassava varieties, thereby contributing to food security, economic development, and the livelihoods of millions who depend on this staple crop. The integration of molecular and systems-level understanding into practical breeding pipelines represents a paradigm shift, moving cassava improvement from conventional phenotype-based selection toward predictive, data-driven, and precision-oriented strategies.

## DATA AVAILABILITY

No new data were generated or analyzed in this study. All data discussed in this review are derived from previously published studies and publicly available sources, which are cited within the article.

## AUTHOR CONTRIBUTIONS

Conceptualization, SPA and IUO; methodology, SPA; software, SPA; validation, CNE, IUO and MJO; formal analysis, SPA; investigation, IUO; resources, SPA; writing—original draft preparation, SPA; writing—review and editing, SPA and MJO; visualization, SPA and IUO; supervision, CNE and IUO. All authors have read and agreed to the published version of the manuscript.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## FUNDING

No external funding was received for this study.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the National Root Crops Research Institute (NRCRI), Umudike, Nigeria, for providing an enabling institutional environment and access to literature, resources, and discussions that supported the preparation of this review. We also appreciate the constructive input and guidance of NRCRI scientific and technical staff, whose expertise and professional advice contributed to shaping the manuscript.

## REFERENCES

1. Howeler RH, Lutaladio N, Thomas G. Save and Grow: Cassava—A Guide to Sustainable Production Intensification. FAO 2013. Available from: <https://www.fao.org/4/i3278e/i3278e.pdf> Accessed on 13 Oct 2025.
2. FAO. FAOSTAT Statistical Database. Food and Agriculture Organization of the United Nations 2021. Available from: <https://www.fao.org/faostat/> Accessed on 10 Sep 2025.
3. Burns AE, Gleadow RM, Cliff J, Zacarias, AM, Cavagnaro TR. Cassava: The drought, war and famine crop in a changing world. *Sustainability*. 2010;2(11):3572-3607. doi: 10.3390/su2113572
4. Jarvis A, Ramirez-Villegas J, Campo BVH, Navarro-Racines C. Is cassava the answer to African climate change adaptation? *Trop Plant Biol.* 2012;5(1):9–29. doi: 10.1007/s12042-012-9096-7
5. Nweke FI, Spencer DSC, Lynam JK. The cassava transformation: Africa's best-kept secret. East Lansing (MI, USA): Michigan State University Press; 2002.
6. FAOSTAT. Food and Agriculture Organization of the United Nations Statistical Database 2020. Available from: <http://www.fao.org/faostat/en/> Accessed on 10 Sep 2025.
7. Ceballos H, Iglesias CA, Pérez JC, Dixon AGO. Cassava breeding: Opportunities and challenges. *Plant Mol Biol.* 2004;56(4):503-16. doi: 10.1007/s11103-004-5010-5

8. Kawuki RS, Ferguson M, Labuschagne M, Herselman L, Kim DJ, Korada R, et al. Variation in starch quality parameters among cassava genotypes in different environments. *J Sci Food Agric.* 2011;91(8):1383-91. doi: 10.1002/jsfa.4319
9. Ceballos H, Kawuki RS, Gracen VE, Yencho GC, Hershey CH. Conventional breeding, marker-assisted selection, genomic selection and inbreeding in clonally propagated crops: a case study for cassava. *Theor Appl Genet.* 2015;128(9):1647-67. doi: 10.1007/s00122-015-2555-4.
10. Fregene M, Bernal A, Duque M, Dixon A, Tohme J, Asiedu R. A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor Appl Genet.* 2001;102(5):703-11. doi: 10.1007/s001220050580
11. Wolfe MD, Del Carpio DP, Alabi O, Ezenwaka LC, Ikeogu UN, Kayondo IS, et al. Prospects for genomic selection in cassava breeding. *Plant Genome.* 2017;10(3). doi:10.3835/plantgenome2017.03.0015
12. Fukuda WMG, Guevara CL, Kawuki R, Ferguson ME. Selected morphological and agronomic descriptors for the characterization of cassava. Ibadan (Oyo, Nigeria): International Institute of Tropical Agriculture; 2010.
13. Prochnik S, Marri PR, Desany B, Rabinowicz PD, Kodira C, Mohiuddin M, et al. The cassava genome: Current progress, future directions. *Trop Plant Biol.* 2012;5(1):88-94. doi: 10.1007/s12042-011-9088-z
14. Bredeson JV, Lyons JB, Prochnik SE, Wu GA, Ha CM, Edsinger-Gonzales E, et al. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat Biotechnol.* 2016;34(5):562-70. doi: 10.1038/nbt.3535
15. Kuon JE, Qi W, Schläpfer P, Hirsch-Hoffmann M, von Bieberstein PR, Patrignani A, et al. Haplotype-resolved genomes of geminivirus-resistant and geminivirus-susceptible African cassava cultivars. *BMC Biol.* 2019;17(1):75. doi: 10.1186/s12915-019-0697-6
16. Rabbi IY, Udoh LI, Wolfe M, Parkes EY, Gedil MA, Dixon A. Genome-Wide Association Mapping of Correlated Traits in Cassava: Dry Matter and Total Carotenoid Content. *Plant Genome.* 2017;10(3). doi: 10.3835/plantgenome2016.09.0094
17. Rabbi IY, Kayondo SI, Bauchet G, Yusuf M, Aghogho CI, Ogunpaimo K, et al. Genome-Wide Association Analysis Reveals New Insights into the Genetic Architecture of Defensive, Agro-Morphological and Quality-Related Traits in Cassava. *Plant Mol Biol.* 2022;109:195-213. doi: 10.1007/s11103-020-01038-3
18. Ma Q, Tong W, Cheng L, Zhang Y, Wang Y, Wang X, et al. Advancing cassava molecular breeding through genome editing: A promising pathway. *Trop Plant.* 2025;4:e004. doi: 10.48130/tp-0024-0046
19. Ceballos H, Hershey C, Becerra López-Lavalle, LA. New Approaches to Cassava Breeding. In: Janick J, editor. *Plant Breeding Reviews.* Hoboken (NJ, USA): John Wiley and Sons, Inc.; 2012. p. 427-504. doi:10.1002/9781118358566.ch6
20. Okogbenin E, Setter TL, Ferguson M, Mutegi R, Ceballos H, Olasanmi B, et al. Phenotypic approaches to drought tolerance in cassava. *Front Physiol.* 2012;3:338. doi: 10.3389/fphys.2012.00338

21. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001;157(4):1819-29. doi: 10.1093/genetics/157.4.1819
22. Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, et al. Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci.* 2017;22(11):961-75. doi: 10.1016/j.tplants.2017.08.011
23. Wolfe MD, Rabbi IY, Egesi C, Hamblin M, Kawuki R, Kulakow P, et al. Genome-Wide Association and Prediction Reveals Genetic Architecture of Cassava Mosaic Disease Resistance and Prospects for Rapid Genetic Improvement. *Plant Genom.* 2016;9(2). doi: 10.3835/plantgenome2015.11.0118
24. Isidro J, Jannink J-L, Akdemir D, Poland J, Heslot N, Sorrells ME. Training set optimization under population structure in genomic selection. *Theor Appl Genet.* 2015;128(1):145-58. doi: 10.1007/s00122-014-2418-4
25. Esuma W, Herselman L, Labuschagne MT, Ramu, P, Lu F, Baguma Y, et al. Genome-wide association mapping of provitamin A carotenoid content in cassava. *Euphytica*. 2016;212:97-110. doi: 10.1007/s10681-016-1772-5
26. Olayinka AF, Dzidzienyo DK, Mbanjo EGN, Offei SK, Tongoona PB, Danquah, EY, et al. Genome-wide association study identifies novel genes for plant architecture and yield traits in cassava (*Manihot esculenta* Crantz). *Front Plant Sci.* 2025;16:1660789. doi: 10.3389/fpls.2025.1660789
27. Odipio J, Alicai T, Ingelbrecht I, Nusinow DA, Bart R, Taylor NJ. Efficient CRISPR/Cas9 genome editing of *phytoene desaturase* in cassava. *Front Plant Sci.* 2017;8:1780. doi: 10.3389/fpls.2017.01780
28. Gomez MA, Lin ZD, Moll T, Luebbert C, Chauhan RD, Hayden L, et al. Simultaneous CRISPR/Cas9-mediated editing of cassava *eIF4E* isoforms *nCBP-1* and *nCBP-2* reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnol J.* 2019;17(2):421-34. doi: 10.1111/pbi.12987
29. Amin A, Zaman W, Park S. Harnessing Multi-Omics and Predictive Modeling for Climate-Resilient Crop Breeding: From Genomes to Fields. *Genes*. 2025;16(7):809. doi: 10.3390/genes16070809
30. Rabbi IY, Hamblin MT, Kumar PL, Gedil MA, Ikpan AS, Jannink J-L, et al. High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res.* 2014;186:87-96. doi: 10.1016/j.virusres.2013.12.028
31. Wang X, Chang L, Tong Z, Wang D, Yin Q, Wang D, et al. Proteomics Profiling Reveals Carbohydrate Metabolic Enzymes and 14-3-3 Proteins Play Important Roles for Starch Accumulation during Cassava Root Tuberization. *Sci Rep.* 2016;6:19643. doi: 10.1038/srep19643
32. Xu X-D, Zhao R-P, Xiao L, Lu L, Gao M, Luo Y-H, et al. Telomere-to-telomere assembly of cassava genome reveals the evolution of cassava and divergence of allelic expression. *Hortic. Res.* 2023;10(11):uhad200. doi: 10.1093/hr/uhad200

33. Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, et al. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*. 2014;9(2):e90346. doi: 10.1371/journal.pone.0090346
34. de Oliveira EJ, de Resende MDV, Santos VdS, Ferreira CF, Oliveira GAF, da Silva MS, et al. Genome-wide selection in cassava. *Euphytica*. 2012;187:263-76, doi: 10.1007/s10681-012-0722-0
35. Delgado A, Hays DB, Bruton RK, Ceballos H, Novo A, Boi E, et al. Ground penetrating radar: a case study for estimating root bulking rate in cassava (*Manihot esculenta* Crantz). *Plant Methods*. 2017;13:65. doi: 10.1186/s13007-017-0216-0
36. Ceballos H, Kulakow P, Hershey C. Cassava breeding: current status, bottlenecks and the potential of biotechnology tools. *Trop Plant Biol*. 2012;5(1):73-87. doi: 10.1007/s12042-012-9094-9
37. Halsey ME, O'Sullivan J, Ceballos H, Fauquet CM, Bellotti AC. Reproductive biology of cassava (*Manihot esculenta* Crantz) and Isolation of Experimental Field Trials. *Crop Sci*. 2008;48(1):49-58. doi: 10.2135/cropsci2007.05.0279
38. Ceballos H, Pérez JC, Joaqui-Barandica O, Lenis JI, Morante N, Calle, F. et al. Cassava Breeding I: The Value of Breeding Value. *Front Plant Sci*. 2016;7:1227. doi: 10.3389/fpls.2016.01227
39. Nassar NMA., Ortiz R. Cassava breeding: Constraints and opportunities. In: Hillocks RJ, Thresh JM, Bellotti AC, editors. *Cassava: Biology, Production and Utilization*. Oxfordshire (UK): CAB International; 2010. p. 149-80.
40. Carvalho LJ, Filho JF, Anderson JV, Figueiredo PW, Chen S. Storage Root of Cassava: Morphological Types, Anatomy, Formation, Growth, Development and Harvest Time. In: Waisundara V, editor. *Cassava*. London (UK): InTech. 2018. doi: 10.5772/intechopen.71347
41. Rüscher D, Corral JM, Carluccio AV, Klemens PAW, Gisel A, Stavolone L, et al. Auxin signaling and vascular cambium formation enable storage metabolism in cassava tuberous roots. *J Exp Bot*. 2021;72(10):3688-703. doi: 10.1093/jxb/erab106
42. El-Sharkawy MA. Cassava biology and physiology. *Plant Mol Biol*. 2004;56(4):481-501. doi: 10.1007/s11103-005-2270-7
43. Suzuki G, Lucob-Agustin N, Kashihara K, Fujii Y, Inukai Y, Gomi K. Rice MEDiator25, OsMED25, is an essential subunit for jasmonate-mediated root development and OsMYC2-mediated leaf senescence. *Plant Sci*. 2021;306:110853. doi: 10.1016/j.plantsci.2021.110853
44. Legg JP, Somado EA, Barker I, Beach L, Ceballos H, Cuéllar WJ, et al. A global alliance declaring war on cassava viruses in Africa. *Food Sec*. 2014;6(2):231-48. doi: 10.1007/s12571-014-0340-x
45. Priyadarshan PM. Genotype-by-Environment Interactions. In: *PLANT BREEDING: Classical to Modern*. Singapore: Springer; 2019. p. 457-72. doi: 10.1007/978-981-13-7095-3\_20
46. Allem AC. The origins and taxonomy of cassava. In: Hillocks RJ, Thresh JM, Bellotti AC, editors. *Cassava: Biology, Production and Utilization*. Oxfordshire (UK): CAB International; 2002. p. 1-16. doi: 10.1079/9780851995243.0001

47. Olsen K, Schaal BA. Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae) and its wild relatives: Further evidence for a southern Amazonian origin of domestication. *Am J Bot.* 2006;93(8):1150-57.
48. Kongsil P, Ceballos H, Siriwan W, Vuttipongchaikij S, Kittipadakul P, Phumichai C, et al. Cassava Breeding and Cultivation Challenges in Thailand: Past, Present, and Future Perspectives. *Plants.* 2024;13(14):1899. doi: 10.3390/plants13141899.
49. Ogbonna AC, Braatz de Andrade LR, Mueller LA, de Oliveira EJ, Bauchet GJ. Comprehensive genotyping of a Brazilian cassava (*Manihot esculenta* Crantz) germplasm bank: Insights into diversification and domestication. *Theor Appl Genet.* 2021;134(5):1343-62. doi: 10.1007/s00122-021-03775-5
50. Ferguson ME, Shah T, Kulakow P, Ceballos H. A global overview of cassava genetic diversity. *PLOS ONE.* 2019;14(11):e0224763. doi: 10.1371/journal.pone.0224763
51. Ramu P, Esuma W, Kawuki R, Rabbi IY, Egesi C, Bredeson JV, et al. Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. *Nat Genet.* 2017;49(6):959-63. doi: 10.1038/ng.3845
52. Nassar NMA. Wild cassava, *Manihot* spp.: Biology and potentialities for genetic improvement. *Genet Mol Biol.* 2000;23(4):1003-14. doi: 10.1590/S1415-47572000000100035
53. Raji AAJ, Fawole I, Gedil M, Dixon, AGO. Genetic differentiation analysis of African cassava (*Manihot esculenta*) landraces and elite germplasm using amplified fragment length polymorphism and simple sequence repeat markers. *Ann Appl Biol.* 2009;155(2):187-99. doi: 10.1111/j.1744-7348.2009.00329.x
54. Mba REC, Stephenson P, Edwards K, Melzer S, Nkumbira J, Gullberg U, et al. Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: Towards an SSR-based molecular genetic map of cassava. *Theor Appl Genet.* 2001;102(1):21-31. doi: 10.1007/s001220051614
55. Elshire R, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE.* 2011;6(5):e19379. doi: 10.1371/journal.pone.0019379
56. Metzker ML. Sequencing technologies—The next generation. *Nat Rev Genet.* 2010;11(1):31-46. doi: 10.1038/nrg2626
57. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nat Rev Genet.* 2011;12(5):363-76. doi: 10.1038/nrg2958
58. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet.* 2011;12(6):443-51. doi: 10.1038/nrg2986
59. Nzuki IW, Katari M, Bredeson JV, Masumba EA, Kapinga F, Salum K, et al. QTL mapping for pest and disease resistance in cassava and coincidence of some QTL with introgression regions derived from *Manihot glaziovii*. *Front Plant Sci.* 2017;8:1168. doi: 10.3389/fpls.2017.01168

60. Ikeogu UN, Akdemir D, Wolfe MD, Okeke UG, Amaefula C, Jannink J-L, et al. Genetic Correlation, Genome-Wide Association and Genomic Prediction of Portable NIRS Predicted Carotenoids in Cassava Roots. *Front Plant Sci.* 2019;10:1570. doi: 10.3389/fpls.2019.01570
61. Udoth LI, Gedil M, Parkes EY, Kulakow P, Adesoye A, Nwuba C, et al. Candidate gene sequencing and validation of SNP markers linked to carotenoid content in cassava (*Manihot esculenta* Crantz). *Mol Breed.* 2017;37(10):123. doi: 10.1007/s11032-017-0718-5
62. Ige AD, Olasanmi B, Bauchet GJ, Kayondo SI, Mbanjo E, Uwugiaren R, et al. Validation of KASP-SNP markers in cassava germplasm for marker-assisted selection of increased carotenoid content and dry matter content. *Front Plant Sci.* 2022;13:1016170. doi: 10.3389/fpls.2022.1016170
63. Codjia ED, Olasanmi B, Ugoji CE, Rabbi IY. SNP-based marker-assisted selection for high provitamin A content in African cassava genetic background. *S Afr J Sci.* 2023;119(11/12): 15115. doi: 10.17159/sajs.2023/15115
64. Prasannakumari V, Heinining Nair AG, Mohan C. Identification of quantitative trait loci (QTLs) conferring dry matter content and starch content in cassava (*Manihot esculenta* Crantz). *Am J BioSci.* 2021;9(1):1-9. doi: 10.11648/j.ajbio.20210901.11
65. Thanyasiriwat T, Sraphet S, Whankaew S, Boonseng O, Lightfoot DA, Tangphatsornruang S, et al. Quantitative trait loci and candidate genes associated with starch pasting viscosity characteristics in cassava (*Manihot esculenta* Crantz). *Plant Biol.* 2013;16(3):699-708. doi: 10.1111/plb.12022
66. Tappiban P, Smith DR, Triwitayakorn K, Bao J. Recent understanding of starch biosynthesis in cassava for quality improvement: A review. *Trends Food Sci Technol.* 2019;83:167-80. doi: 10.1016/j.tifs.2018.11.019
67. Whankaew S, Poopear S, Kanjanawattanawong S, Tangphatsornruang S, Boonseng O, Lightfoot DA, et al. A genome scan for quantitative trait loci affecting cyanogenic potential of cassava root in an outbred population. *BMC Genom.* 2011;12:266. doi: 10.1186/1471-2164-12-266
68. Adjepong-Danquah J, Gracen VE, Offei SK, Asante IK, Manu-Aduening JA. Genetic variability in storage root bulking of cassava genotypes under irrigation and no irrigation. *Agric Food Secur.* 2016;5:1-12. doi: 10.1186/s40066-016-0055-7
69. Okogbenin E, Fregene M. Genetic analysis and QTL mapping of early root bulking in an F1 population of non-inbred parents in cassava (*Manihot esculenta* Crantz). *Theor Appl Genet.* 2002;106:58-66. doi: 10.1007/s00122-002-1068-0
70. Gorjanc G, Bijma P, Hickey JM. Reliability of pedigree-based and genomic evaluations in selected populations. *Genet Sel Evol.* 2015;47:65. doi: 10.1186/s12711-015-0145-1
71. Abah SP, Mbe JO, Dzidzienyo DK, Njokum D, Onyeka J, Danquah EY, et al. Determination of genomic regions associated with early storage root formation and bulking in cassava. *Front Plant Sci.* 2024;15:1391452. doi: 10.3389/fpls.2024.1391452

72. Luo S, Ma Q, Zhong Y, Jing J, Wei Z, Zhou W, et al. Editing of the starch branching enzyme gene SBE2 generates high-amylose storage roots in cassava. *Plant Biotechnol J.* 2022;18(8):1835-47. doi: 10.1007/s11103-021-01215-y
73. Lyons JB, Bredeson V, Mansfeld BN, Bauchet GJ, Berry , Boyher A, et al. Current status and impending progress for cassava structural genomics. *Plant Mol Biol.* 2022;109(3):177-91. doi: 10.1007/s11103-020-01104-w
74. Jørgensen K, Bak S, Busk PK, Sørensen C, Olsen CE, Puonti-Kaerlas J, et al. Cassava plants with a depleted cyanogenic glucoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. *Plant Physiol.* 2005;139(1):363-74. doi: 10.1104/pp.105.065904
75. Welsch R, Arango J, Bär C, Salazar B, Al-Babili S, Beltrán J, et al. Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *Plant Cell.* 2010;22(10):3348-56. doi: 10.1105/tpc.110.077560
76. Udoth LI, Agogbu JU, Keyagha ER, Nkanga II. Carotenoids in cassava (*Manihot esculenta* Crantz): Genes regulating  $\beta$ -carotene levels and implications for biofortification. In: Martínez-Espinosa RM, editor. Carotenoids—New Perspectives and Applications. London (UK): IntechOpen; 2022 doi: 10.5772/intechopen.105210
77. Orek C. A review of drought-stress responsive genes and their applications for drought stress tolerance in cassava (*Manihot esculenta* Crantz). *Discov Biotechnol.* 2024;1:5. doi: 10.1007/s44340-024-00006-7
78. Koundinya AVV, Muthusamy SK, Ajeesh BR, Mohan C, Sreekumar J, Pulapet S, et al. Elucidation of novel drought-responsive genes from tuber transcriptome of cassava under water deficit stress. *S Afr J Bot.* 2024;169:255-67. doi: 10.1016/j.sajb.2024.04.029
79. Allie F, Pierce EJ, Okoniewski MJ, Rey C. Transcriptional analysis of South African cassava mosaic virus-infected susceptible and tolerant landraces of cassava highlights differences in resistance, basal defense and cell wall associated genes during infection. *BMC Genom.* 2014;15:1006. doi: 10.1186/1471-2164-15-1006
80. Kung'u MW, Rey C, Nyaboga EN. Identification and Characterization of Nbs-Lrr Genes in Cassava and Their Expression Profiles in Susceptible and Resistant Cassava Varieties in Response to East African Cassava Mosaic Virus Infection. Available from: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5223198](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5223198). Accessed 18 Sep 2025.
81. Freeborough W, Gentle N, Rey, MEC. WRKY Transcription Factors in Cassava Contribute to Regulation of Tolerance and Susceptibility to Cassava Mosaic Disease through Stress Responses. *Viruses.* 2021;13(9):1820. doi: 10.3390/v13091820

82. Wahengbam ED, Devi CP, Sharma SK, Roy SS, Maibam A, Dasgupta M, et al. Reactive oxygen species turnover, phenolics metabolism, and some key gene expressions modulate postharvest physiological deterioration in cassava tubers. *Front Microbiol.* 2023;14:1148464. doi: 10.3389/fmicb.2023.1148464
83. Djabou ASM, Carvalho LJCB, Li QX, Niemenak N, Chen S. Cassava postharvest physiological deterioration: A complex phenomenon involving calcium signaling, reactive oxygen species and programmed cell death. *Acta Physiol Plant.* 2017;39:91. doi: 10.1007/s11738-017-2382-0
84. Montesinos-López OA, Montesinos-López A, Mosqueda-González BA, Delgado-Enciso I, Chavira-Flores M, Crossa J, et al. Genomic prediction powered by multi-omics data. *Front Genet.* 2025;16:1636438. doi: 10.3389/fgene.2025.1636438
85. Landi M, Carluccio AV, Shah T, Niazi A, Stavolone L, Falquet L, et al. Genome-wide comparison reveals large structural variants in cassava landraces. *BMC Genom.* 2025;26:362. doi: 10.1186/s12864-025-11523-y
86. Santos CCd, Andrade LRBd, Carmo CDd, Oliveira EJd. Development of cassava core collections based on morphological and agronomic traits and SNPs markers. *Front Plant Sci.* 2023;14:1250205. doi: 10.3389/fpls.2023.1250205
87. Morales N, Ogbonna AC, Ellerbrock BJ, Bauchet GJ, Tantikanjana T, Tecle IY, et al. Breedbase: a digital ecosystem for modern plant breeding. *G3: Genes Genomes Genet.* 2022;12(7):jkac078. doi: 10.1093/g3journal/jkac078
88. Patil BL, Fauquet CM. Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol Plant Pathol.* 2009;10(5):685-701. doi: 10.1111/j.1364-3703.2009.00559.x
89. Ospina JA, López-Alvarez D, Gimode W, Wenzl P, Carvajal-Yepes, M. Genome-wide association study of cassava brown streak disease resistance in cassava germplasm conserved in South America. *Sci Rep.* 2024;14:23141. doi: 10.1038/s41598-024-74161-6
90. Santos CSd, Sousa MB, Brito AC, de Oliveira LA, Carvalho CWP, de Oliveira EJ. Genome-wide association study of cassava starch paste properties. *PLOS ONE.* 2022;17(1):e0262888. doi: 10.1371/journal.pone.0262888
91. Sampaio Filho JS, Olivoto T, Campos Mds, Oliveira EJd. Multi-trait selection in multi-environments for performance and stability in cassava genotypes. *Front Plant Sci.* 2023;14:1282221. doi: 10.3389/fpls.2023.1282221
92. Jarquín D, Crossa J, Lacaze X, du Cheyron P, Daucourt J, Lorgeou, J. et al. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor Appl Genet.* 2014;127(3):595-607. doi: 10.1007/s00122-013-2243-1
93. Ozimati AA, Esuma W, Manze F, Iragaba P, Kanaabi M, Ano CU, et al. Utility of Ugandan genomic selection cassava breeding populations for prediction of cassava viral disease resistance and yield in West African clones. *Front Plant Sci.* 2022;13:1018156. doi: 10.3389/fpls.2022.1018156
94. Andrade LRBd, Sousa MBe, Oliveira EJ, Resende MDVd, Azevedo CF. Cassava yield traits predicted by genomic selection methods. *PLoS ONE.* 2019;14(11):e0224920. doi: 10.1371/journal.pone.0224920

95. Filho JSS, Oliveira ICM, Pastina MM, Campos MdS, de Oliveira EJ. Genotype x environment interaction in cassava multi-environment trials via analytic factor. *PLoS ONE*. 2024;19(12):e0315370. doi: 10.1371/journal.pone.0315370
96. Zhang R, Zhang C, Yu C, Dong J, Hu J. Integration of multi-omics technologies for crop improvement: Status and prospects. *Front Bioinform*. 2022;2:1027457. doi: 10.3389/fbinf.2022.1027457
97. Obata T, Fernie AR. The use of metabolomics to dissect plant responses to abiotic stresses. *Cell Mol Life Sci*. 2012;69(19):3225-43. doi: 10.1007/s00018-012-1091-5
98. Montagnac JA, Davis CR, Tanumihardjo SA. Nutritional Value of Cassava for Use as a Staple Food and Recent Advances for Improvement. *Compr Rev Food Sci Food Saf*. 2009;8(3):181-94. doi: 10.1111/j.1541-4337.2009.00077.x.
99. Younas MU, Zuo S, Qasim M, Ahmad I, Feng Z, Korai SK., et al. Multi-omics approaches in plant biology: decoding agronomic traits for sustainable agriculture. *Plant Stress*. 2025;18:101118, ISSN 2667-064X, doi: 10.1016/j.stress.2025.101118
100. Perez-Fons L, Ovalle TM, Maruthi MN, Colvin J, Lopez-Lavalle LAB, Fraser PD. The metabotyping of an East African cassava diversity panel: a core collection for developing biotic stress tolerance in cassava. *PLoS ONE*. 2020;15(11):e0242245. doi: 10.1371/journal.pone.0242245
101. Uarrota VG, Maraschin M. Metabolomic, enzymatic, and histochemical analyzes of cassava roots during postharvest physiological deterioration. *BMC Res Notes*. 2015;8:648. doi: 10.1186/s13104-015-1580-3
102. Singh N. Starch in Food: Structure, Function and Applications. *Int J Food Sci Technol*. 2006;41(1):108-9, doi: 10.1111/j.1365-2621.2005.00966.x
103. Rosado-Souza L, David LC, Drapal M, Fraser PD, Hofmann J, Klemens PAW, et al. Cassava Metabolomics and Starch Quality. *Curr Protoc Plant Biol*. 2019;4(4):e20102. doi: 10.1002/cppb.20102.
104. Xiao L, Cheng D, Ou W, Chen X, Rabbi IY, Wang W, et al. Advancements and strategies of genetic improvement in cassava (*Manihot esculenta* Crantz): from conventional to genomic approaches. *Hortic Res*. 2024;12(3):uhae341. doi: 10.1093/hr/uhae341
105. Maziya-Dixon B, Dixon AGO, Adebawale AA. Targeting different end uses of cassava: Genotypic variations for cyanogenic potentials and pasting properties. *Int J Food Sci Technol*. 2007;42:969-76. doi: 10.1111/j.1365-2621.2006.01319.x
106. Ding Z, Fu L, Wang B, Ye J, Ou W, Yan Y, et al. Metabolic GWAS-based dissection of genetic basis underlying nutrient quality variation and domestication of cassava storage root. *Genome Biol*. 2023;24:289. doi: 10.1186/s13059-023-03137-y
107. Amelework AB, Bairu MW, Marx R, Laing M, Venter SL. Genotype x Environment Interaction and Stability Analysis of Selected Cassava Cultivars in South Africa. *Plants*. 2023;12(13):2490. doi: 10.3390/plants12132490
108. Anjanappa RB, Gruissem W. Current progress and challenges in crop genetic transformation. *J Plant Physiol*. 2021;261:153411. doi: 10.1016/j.jplph.2021.153411

109. Chaowongdee S, Vannatim N, Malichan S, Kuncharoen N, Tongyoo P, Siriwan W. Comparative transcriptomics analysis reveals defense mechanisms of *Manihot esculenta* Crantz against Sri Lanka Cassava Mosaic Virus. *BMC Genom.* 2024;25(1):436. doi: 10.1186/s12864-024-10315-0
110. Orek, C. A review of the functions of transcription factors and related genes involved in cassava (*Manihot esculenta* Crantz) response to drought stress. *Trop Plants.* 2023;2:14. doi: 10.48130/TP-2023-0014
111. Zhou Z, Liu J, Liu J. Application of Weighted Gene Co-Expression Network Analysis to Metabolomic Data from an ApoA-I Knockout Mouse Model. *Molecules.* 2024;29(3):694. doi: 10.3390/molecules29030694
112. Naconsie M, Lertpanyasampatha M, Viboonjun U, Netrphan S, Kuwano M, Ogasawara N, et al. Cassava root membrane proteome reveals activities during storage root maturation. *J Plant Res.* 2016;129(1):51-65. doi:10.1007/s10265-015-0761-4
113. Amelework AB, Bairu MW. Advances in Genetic Analysis and Breeding of Cassava (*Manihot esculenta* Crantz): A Review. *Plants.* 2022;11(12):1617. doi: 10.3390/plants11121617
114. Hu W, Ji C, Liang Z, Ye J, Ou W, Ding Z, et al. Resequencing of 388 cassava accessions identifies valuable loci and selection for variation in heterozygosity. *Genome Biol.* 2021;22:316. doi: 10.1186/s13059-021-02524-7
115. CassavaBase: A multi-crop database for the management of genomic, genetic, and breeding data. Available from: <https://cassavabase.org/>. Accessed 20 October 2025.
116. Ahmad S, Shah MH, Raghbir MG, Belwal V, Ali S, Khan R, et al. Technical Advancements in Plant High-Throughput Phenotyping and Phenomics. In: Plant High-Throughput Phenotyping and Functional Phenomics. Boca Raton (FL, USA): CRC Press; 2025. p. 14.

How to cite this article:

Abah SP, Okpani MJ, Okwuonu IU, Egesi CN. Genomics-Assisted Cassava Breeding: A Comprehensive Review of Advances, Tools, and Future Prospects. *Crop Breed Genet Genom.* 2026;8(1):e20260002. <https://doi.org/10.20900/cbgg20260002>.