RGSA – Revista de Gestão Social e Ambiental
ISSN: 1981-982X
Submission date: 07/19/2024
Acceptance date: 09/20/2024
DOI: https://doi.org/10.24857/rgsa.v18n10-313
Organization: Interinstitutional Scientific Committee
Chief Editor: Ana Carolina Messias de Souza Ferreira da Costa
Assessment: Double Blind Review pelo SEER/OJS

GENOMIC-WIDE ASSOCIATION STUDIES FOR RESISTANCE TO ANTHRACNOSE IN A PSEUDO-F_2 MANGO POPULATION

Carlos Antônio Fernandes Santos[1]
Deisy Aiane Lima de Aquino[2]
Diógenes da Cruz Batista[3]

ABSTRACT

Goal: This study aimed to apply genome-wide association studies (GWAS) to identify single nucleotide polymorphisms (SNPs) associated with resistance to mango anthracnose.

Theoretical Framework: The statistical methods of GWAS aim to identify associations among genotypes with phenotypes for marker-assisted selection (MAS) or genome dissection of important traits.

Method: Ninety-four plants of a 'Haden' × 'Tommy Atkins' pseudo-F_2 population were genotyped with 705 SNPs and phenotyped for resistance to *Colletotrichum siamense*, anthracnose causal agent, by spraying young leaves, three times at different periods, with 10^3 conidia/mL. Qualitative GWAS, with Plink Fisher's exact test and Cochran-Armitage trend genotypic test, and quantitative GWAS, with Tassel general (GLM) and mixed linear models (MLM) were applied. F2 plants with symptoms mean absence >75% were considered moderately resistant to the fungus.

Results and discussion: Qualitative and quantitative GWAS methods combined enable the identification of two consensus SNP loci controlling 24.76% of total variance to mango anthracnose resistance. Other four SNPs, identified, exclusively in quantitative or qualitative analysis, explained 28.55% of total variance to mango anthracnose. These SNPs loci are promising candidate for further PCR primers design in order to apply MAS in mango.

Research implications: This study facilitates the identification of consensus chromosomal regions associated with mango anthracnose resistance. The application of MAS in mango breeding will accelerate the development of new cultivars.

Originality/Value: To our knowledge, this a pioneering GWAS study applied to anthracnose, which is the most significant and widespread fungal disease affecting mango fruit globally.

Keywords: *Mangifera indica*, GWAS, SNP, Plink, Tassel.

ESTUDOS DE ASSOCIAÇÃO GENÔMICA AMPLA PARA RESISTÊNCIA À ANTRACNOSE EM POPULAÇÃO PSEUDO-F2 DE MANGA

RESUMO

Objetivo: O objetivo foi aplicar estudos de associação genômica ampla (GWAS) para identificar polimorfismo de nucleotídeo único (SNP) associado à antracnose de manga.

Referencial Teórico: métodos estatísticos de GWAS visam identificar associações de genótipos com fenótipos para seleção assistida por marcadores (MAS) ou dissecção do genoma para características importantes.

[1] Embrapa Semiárido, Petrolina, PE, Brasil. E-mail: carlos-fernandes.santos@embrapa.br
Orcid https://orcid.org/0000-0002-6932-6805
[2] Universidade Estadual de Feira de Santana, Feira de Santana, BA, Brasil. E-mail: deisylima_17@hotmail.com
Orcid https://orcid.org/0000-0001-8897-320X
[3] Embrapa Semiárido Embrapa Semiárido, Petrolina, PE, Brasil. E-mail: diogenes.batista@embrapa.br
Orcid: https://orcid.org/0000-0003-0245-9256

Método: Noventa e quatro plantas da população pseudo-F2 'Haden' × 'Tommy Atkins' foram genotipadas com 705 SNPs e fenotipadas para *Colletotrichum siamense*, agente causal da antracnose, pulverizando folhas jovens, três vezes em diferentes períodos, com 10^3 conídios/mL. Foram aplicados GWAS qualitativos (Plink), com teste exato de Fisher e teste genotípico de 'Cochran-Armitage trend', e GWAS quantitativos (Tassel), com modelos lineares geral e misto. Plantas F2 com ausência média de sintomas >75% foram consideradas moderadamente resistentes ao fungo.

Resultados e discussão: a aplicação conjunta de métodos qualitativos e quantitativos de GWAS permitiu identificar dois loci SNP consenso, controlando 24,76% da variância total para resistência à antracnose. Outros quatro SNPs, identificados exclusivamente na análise quantitativa ou qualitativa, explicaram 28,55% da variância total para a antracnose. Esses loci SNPs são candidatos para desenhos de primers de PCR, para aplicar MAS em manga.

Implicações de pesquisa: A presente abordagem permitiu identificar regiões cromossômicas de consenso relacionadas à antracnose da manga. A aplicação do MAS na manga irá acelerar o desenvolvimento de novos cultivares.

Originalidade/Valor: estudo de GWAS pioneiro aplicado à antracnose, a doença fúngica mais importante e difundida afetando a manga em todo o mundo.

Palavras-chave: *Mangifera indica*, GWAS, SNP, Plink, Tassel.

ESTUDIOS DE ASOCIACIÓN GENÓMICA PARA LA RESISTENCIA A LA ANTRACNOSIS EN POBLACIONES DE MANGO PSEUDO-F2

RESUMEN

Objetivo: El objetivo fue aplicar estudios de asociación de todo el genoma (GWAS) para identificar el polimorfismo de un solo nucleótido (SNP) asociado con la antracnosis del mango.

Marco teórico: Los métodos estadísticos GWAS tienen como objetivo identificar asociaciones de genotipos con fenotipos para la selección asistida por marcadores (MAS) o la disección del genoma para rasgos importantes.

Método: Se genotiparon noventa y cuatro plantas de la población pseudo-F2 'Haden' × 'Tommy Atkins' con 705 SNP y se fenotiparon para *Colletotrichum siamense*, el agente causal de la antracnosis, rociando hojas jóvenes, tres veces en diferentes períodos, con 103 conidias/mL. Se aplicaron GWAS cualitativos (Plink), con prueba exacta de Fisher y prueba genotípica 'tendencia de Cochran-Armitage', y GWAS cuantitativos (Tassel), con modelos lineales generales y mixtos. Las plantas F2 con una ausencia promedio de síntomas >75% se consideraron moderadamente resistentes al hongo.

Resultados y discusión: la aplicación conjunta de métodos GWAS cualitativos y cuantitativos permitió la identificación de dos loci de SNP de consenso, controlando el 24,76% de la varianza total para la resistencia a la antracnosis. Otros cuatro SNP, identificados exclusivamente en el análisis cuantitativo o cualitativo, explicaron el 28,55% de la varianza total de la antracnosis. Estos loci SNP son candidatos para diseños de cebadores de PCR para aplicar MAS en mango.

Implicaciones de la investigación: El presente enfoque nos permitió identificar regiones cromosómicas de consenso relacionadas con la antracnosis del mango. La aplicación de MAS al mango acelerará el desarrollo de nuevos cultivares.

Originalidad/Valor: Estudio pionero GWAS aplicado a la antracnosis, la enfermedad fúngica más importante y extendida que afecta al mango en todo el mundo.

Palabras clave: *Mangifera indica*, GWAS, SNP, Plink, Borla.

1 INTRODUCTION

Mango (*Mangifera indica* L., 2n = 40 chromosomes) is a fruit tree that originated in Eastern India, from Assam to Myanmar, and later spread to Southeast Asia, Africa, and the Americas (Duval *et al*., 2005). In 2022, the largest mango exporters were Mexico (447,502 tons), Thailand (337,263 tons), Peru (241,182 tons), and Brazil (231,874 tons), while the largest importers were the United States (574,782 tons), the Netherlands (278,253 tons), the United Arab Emirates (89,455 tons), and Germany (87,397 tons) (FAOSTAT, 2024). The micro-regions of Juazeiro, BA, and Petrolina, PE, account for more than 86% of Brazil's exports to Europe and the USA (Brasil, 2022). Mango fruits have an attractive visual appearance, and their popularity is growing due to their high nutraceutical and pharmaceutical value, including the polyphenol mangiferin, the most abundant bioactive compound (Wang *et al*., 2020).

Anthracnose is a disease that affects mango production in most growing regions. Symptoms appear on plant parts such as young leaves, causing deformation, twisting, necrosis, and the formation of crusts with ruptures in the affected area. In fruits, the pathogen remains latent and typically manifests after harvest, during the ripening process, as dark, slightly sunken, usually round spots (Terao, Batista & Ribeiro, 2016). In northeastern Brazil, the fungus *Colletotrichum siamense* (=*C. dianesei*) is the most frequently associated with mango damage, although other *Colletotrichum* species also infect the crop to a lesser extent (Lima *et al*., 2013). Anthracnose causes losses ranging from 2% to 60% in mango production across various countries (Lima *et al*., 2013; Tucho, Lemessa & Berecha, 2014).

According to Aruz (2000), mango anthracnose is more virulent under high temperatures and humidity, with an incidence of nearly 100% in fruits produced in very humid conditions. Additionally, Lima *et al*. (2015) observed that lesions caused by *Colletotrichum* species were larger at temperatures between 25°C and 30°C compared to those at 10°C.

The most common method of disease control is the use of fungicides, applied both before fruit development (pre-harvest) and post-harvest, often in conjunction with heat treatment (Ploetz, 2010). However, fungicides with higher acceptability are not always effective and are often not approved in countries like the United States (Ploetz, 2010). The use of unregistered products may result in the loss of export licenses (Anvisa, 2016).

Preventive applications of pre- and post-harvest fungicides do not always guarantee success due to the complex management of the disease and its various epidemiological factors (Aruz, 2000). Genetic control could significantly reduce production costs while improving post-

harvest quality and fruit shelf life, making genetic resistance a more economical, sustainable, and environmentally safe approach (Gupta, Singh & Singh, 2015).

The chromosome-level sequencing and assembly of the mango genome, cv. Alphonso (Wang *et al*., 2020), enables trait association studies and fine mapping of this species. Dautt-Castro *et al*. (2019) reported the genome-wide identification of mango polygalacturonases based on the Tommy Atkins mango genome. Santos *et al*. (2023) identified genomic regions associated with resistance to *Lasiodiplodia theobromae* and *Neofusicoccum parvum* through genome-wide association studies (GWAS) in a pseudo-F2 mango population. Ma *et al*. (2024) identified several candidate loci and dominant genotypes associated with key mango traits by applying GWAS with a large number of SNPs.

2 THEORETICAL FRAMEWORK

GWAS aim to identify associations among genotypes and phenotypes by testing for differences in the allele frequency of genetic variants among individuals who are ancestrally similar but differ phenotypically. The most commonly studied genetic variants in GWAS are single-nucleotide polymorphisms (SNPs) (Uffelmann *et al*., 2021). An SNP is a variation in a single nucleotide (i.e., A, C, G, or T) at a specific position in the genome, typically existing in two forms (e.g., A vs. T), known as alleles (Marees *et al*., 2018). GWAS usually require a large population, often more than 1,000 individuals, and thousands of SNPs (Marees *et al*., 2018). Despite the attractiveness of GWAS, initial studies indicated low heritability values (heritability is the proportion of trait variation attributable to genetic alleles) for several traits (Maher, 2008).

Quantitative Trait Loci (QTL) analysis encompasses various statistical procedures for studying traits with continuous inheritance, such as yield. It involves associating genetic linkage maps (pseudo-chromosomes) with phenotypic data, usually measured in the field. Consistent successes in QTL mapping have been achieved mainly in species where inbred strains or lines can be developed; in humans, QTL mapping has shown limited success (Majumder & Ghosh, 2005).

GWAS and QTL mapping are often complementary, and when conducted together, they offset each other's limitations (Korte & Farlow, 2013). The ultimate goal of both methodologies is the application of marker-assisted selection (MAS) to accelerate the development of new plant cultivars.

In the same mango pseudo-F2 population of the present study, linkage map development in the Haden × Tommy Atkins cross resulted in up to 23 groups, making QTL analysis

impossible (Santos *et al*., 2023). To overcome these limitations, Santos *et al*. (2023) applied qualitative and quantitative genome association methods to identify consensus chromosomal regions controlling diseases.

From this perspective, the present study aimed to apply genome-wide association studies (GWAS) to identify single-nucleotide polymorphisms (SNPs) associated with *Colletotrichum siamense*, supporting the development of marker-assisted selection (MAS) for the improvement of new mango cultivars.

3 METHODOLOGY

3.1 PLANT MATERIAL AND DNA EXTRACTION

Ninety-four plants from a 'Haden' × 'Tommy Atkins' pseudo-F2 population were obtained by identifying two isolated 'Haden' plants surrounded by hundreds of 'Tommy Atkins' plants in a commercial orchard located in Petrolina, PE, Brazil. The progenies were developed in a nursery, and after reaching 50 cm in height, they were established in the field, in 2002, at the Mandacaru Experimental Station of Embrapa, Juazeiro, BA, Brazil. Standard crop management practices widely adopted for mango in the region were applied, including early floral induction using the plant growth regulator paclobutrazol (PBZ) and potassium nitrate applications, combined with pruning and irrigation. Selfed and off-type plants were identified by SNP markers and excluded from genetic analyses. According to Lima de Aquino *et al*. (2021), 'Haden' is resistant, while 'Tommy Atkins' is moderately susceptible to the fungus.

3.2 PHENOTYPING FOR COLLETOTRICHUM SIAMENSE

Colletotrichum siamense isolates were provided by the mycology laboratory at the Federal Rural University of Pernambuco (UFRPE), Recife, PE. The fungus was cultured for 15 days on water-agar in Petri dishes, and conidia were extracted by removing the pycnidia from the colonies. Pycnidia were ground with a pestle in a mortar containing sterilized distilled water. The suspension was filtered, and conidia were counted using a Neubauer chamber, adjusted to a concentration of 10^3 conidia/mL.

The inoculum suspension (10^3 conidia/mL) was gently sprayed onto healthy young branches with young leaves using a hand sprayer until runoff. After inoculation, the branches

were placed in a wet chamber made of plastic bags moistened with sterilized distilled water for 48 hours. To prevent leaf and branch damage from solar radiation, the plastic bags were covered with paper bags. Anthracnose symptoms were evaluated five days after removing the wet chamber. This procedure was repeated three times at different periods on each 'Haden' × 'Tommy Atkins' progeny, with symptom evaluations performed by an experienced rater.

The method for evaluating disease symptoms was adapted from the percentage infection method described by Lima de Aquino *et al*. (2021): plants without symptoms were calculated as (Number of branches without symptoms / Total number of inoculated branches) × 100. Plants with no symptoms and a mean greater than 75% were considered moderately resistant to *C. siamense*.

3.3 GENOTYPING, PHYSICAL MAPPING, AND STATISTICAL ANALYSIS

Pseudo-F2 plants were genotyped with 705 SNPs, as described by Kuhn *et al*. (2017). For qualitative genomic associations, based on the presence or absence of disease symptoms, allelic and Fisher's exact tests were conducted, along with genotypic association using the Cochran-Armitage trend test. These analyses were implemented using the DOS Plink software (Purcell *et al*., 2007). Plants were coded as 1 for resistant or controls (absence of symptoms >75%) and 2 for susceptible or cases (presence of symptoms <75%) to the fungus.

Quantitative genomic associations were performed using the general linear model (GLM) and the mixed linear model (MLM) in Tassel 5.2.65 software (Bradbury *et al*., 2007). In the GLM analysis, genotypic and phenotypic data, both with and without principal component analysis (PCA), were considered. In the MLM analysis, an identical by state (IBS) matrix, estimated using Plink 1.9 (Purcell *et al*., 2007), was included. For PCA estimation, the accumulation of 50% of total variation in the first five eigenvectors was considered (Bradbury *et al*., 2007). Percentage data were transformed to arcsine for GLM and MLM analyses.

SNP sequences, published by Kuhn *et al*. (2017), were aligned to the *Mangifera indica* genome (Wang *et al*., 2020) available on NCBI. Loci were assigned to mango chromosomes based on high e-values and bit scores. SNPs were ordered by chromosome in an Excel spreadsheet.

4 RESULTS AND DISCUSSION

Nineteen pseudo-F2 plants (A1-103, A1-112, A1-136, A1-15, A1-33, A1-36, A1-38, A1-43, A1-75, A1-82, A1-83, A1-85, A2-40, A2-54, A2-56, A2-62, A2-65, A2-70, and A2-96) exhibited moderate resistance (mean >75%) to *Colletotrichum siamense* across three field inoculation trials. The anthracnose symptoms observed in the inoculated plants included leaf deformities, twisting, necrosis, and the formation of crusts, characteristic of the disease (Aquino *et al*., 2021). Pruning was applied to prevent the disease from spreading throughout the plants, which could otherwise lead to plant death and loss of the pseudo-F2 segregating population.

The average number of SNPs mapped to the 'Alphonso' mango cultivar genome (Wang *et al*., 2020) per chromosome was 36.6, ranging from 18 on chromosome 18 to 70 on chromosome 5. Significant gaps in SNP coverage were observed across all chromosomes, with coverage extending up to 6,711,197 bp on chromosome 6 of 'Alphonso.' Gaps were detected at either the beginning or end of chromosomes 2, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 17, and 19, while gaps were observed in the middle of other chromosomes (data not shown).

4.1 ANTHRACNOSE QUALITATIVE GENOMIC ASSOCIATION - ALLELIC AND GENOTYPIC TESTS

The number of SNP loci identified in Fisher's exact test varied based on different resistance levels to *C. siamense* (Table 1). Notably, loci Mi_0092 and Mi_0169 on chromosome 16 were significantly associated with resistance levels of >75%, >80%, and >85%. Additionally, the Mango_rep_c758 locus on chromosome 1 and Mi_0542 locus on chromosome 9 were significantly associated with resistance levels of >75% and >80%, but not at >85% (Table 1).

Table 1

Fisher's exact test and Cochran-Armitage trend genotypic test for SNP loci with statistical associations (P-value<0.009) with Colletotrichum resistance at various resistance levels in the 'Haden' × 'Tommy Atkins' mango pseudo-F_2 population.

SNP	Chromosome	Position (bp)	*P-value*
Fisher's exact test - resistance level > 75%			
Mi_0092	16	4787365	0.0007075
Mi_0169	16	4815976	0.0007075
Mango_rep_c758	1	4082294	0.005272
Mi_0542	9	1528958	0.007353
Mi_0213	14	688814	0.008373
Mi_0317	5	19641916	0.008385
Fisher's exact test - resistance level > 80%			
Mi_0542	9	1528958	0.00153
Mi_0460	9	1422430	0.003904
Mi_0092	16	4787365	0.004138
Mi_0169	16	4815976	0.004138
Mango_rep_c758	1	4082294	0.008915
Fisher's exact test - resistance level > 85%			
Mi_0092	16	4787365	0.003552
Mi_0169	16	4815976	0.003552
Cochran-Armitage trend genotypic test – resistance level > 75%			
Mi_0092	16	4787365	0.0002861
Mi_0169	16	4815976	0.0002861
Mi_0317	5	19641916	0.0009959
Mi_0213	14	688814	0.001253
Mi_0187	9	8122380	0.001963
Mango_rep_c758	1	4082294	0.002759
Mi_0254	4	9321622	0.003371
Mi_0239	2	18380580	0.007526
Mi_0542	9	1528958	0.007944
Cochran-Armitage trend genotypic test – resistance level > 80%			
Mi_0542	9	1528958	0.001772
Mi_0092	16	4787365	0.00207
Mi_0169	16	4815976	0.00207
Mi_0317	5	19641916	0.002331
Mi_0460	9	1422430	0.002698
Mi_0213	14	688814	0.002794
Mango_rep_c758	1	4082294	0.005459
Cochran-Armitage trend genotypic test – resistance level > 85%			
Mi_0092	16	4787365	0.002052
Mi_0169	16	4815976	0.002052
Mango_rep_c758	1	4082294	0.00719
Contig_6876_G118T	13	8943807	0.008609

The number of SNP loci identified in the Cochran-Armitage trend genotypic test also varied with different resistance levels to *C. siamense* (Table 1). The Mi_0092 and Mi_0169 loci on chromosome 16 and the Mango_rep_c758 locus on chromosome 1 were significantly associated with resistance levels of >75%, >80%, and >85%. Additionally, the Mi_0542, Mi_0317, and Mi_0213 loci, located on chromosomes 9, 5, and 14, respectively, were significantly associated with resistance levels of >75% and >80%, but not at >85%, in the Cochran-Armitage test (Table 1).

An increased number of SNP loci were identified in the Cochran-Armitage trend test compared to Fisher's exact test across all resistance levels (Table 1). This increase can be attributed to the differences in the assumptions underlying these tests: the Cochran-Armitage trend test analyzes individual samples without Hardy-Weinberg equilibrium (HWE), while Fisher's exact test analyzes allele samples in HWE. Similar findings were reported by Santos *et al*. (2023) in a GWAS study on dieback fungi in mango.

A reduced number of SNP loci were associated with increasing resistance levels, from >75% to >85%, for both allelic and genotypic tests (Table 1). This reduction is partly due to the smaller number of progenies with resistance levels >75% (19 progenies), compared to 74 progenies with resistance levels <75%, indicating that susceptibility is dominant in a cross between a resistant and a moderately susceptible parent. Aquino *et al*. (2021) found only eight out of 82 progenies with resistance levels >89% to *C. siamense* in the same F2 population, which corroborates our findings.

Loci Mi_0092 and Mi_0169, located close together on chromosome 16, were significantly associated with resistance levels of >75%, >80%, and >85% in both Fisher's exact test and the Cochran-Armitage test (Table 1). These loci are strong candidates for major resistance loci to anthracnose in mango.

4.2 ANTHRACNOSE QUANTITATIVE GENOMIC ASSOCIATION - GENERAL AND MIXED LINEAR MODELS

The number of SNP loci identified in GLM analysis, with and without principal component stratification (PCS), varied except for locus Mi_0187, which showed different significance levels (P-values of 0.000605 and 0.009057, respectively) (Table 2, Fig. 1). Only one SNP locus, Mi_0253, was common in MLM analysis, both with and without PCS. The Mi_0187 locus was significantly associated with resistance in both GLM and MLM analyses, except in MLM with PCS (Table 2).

Table 2

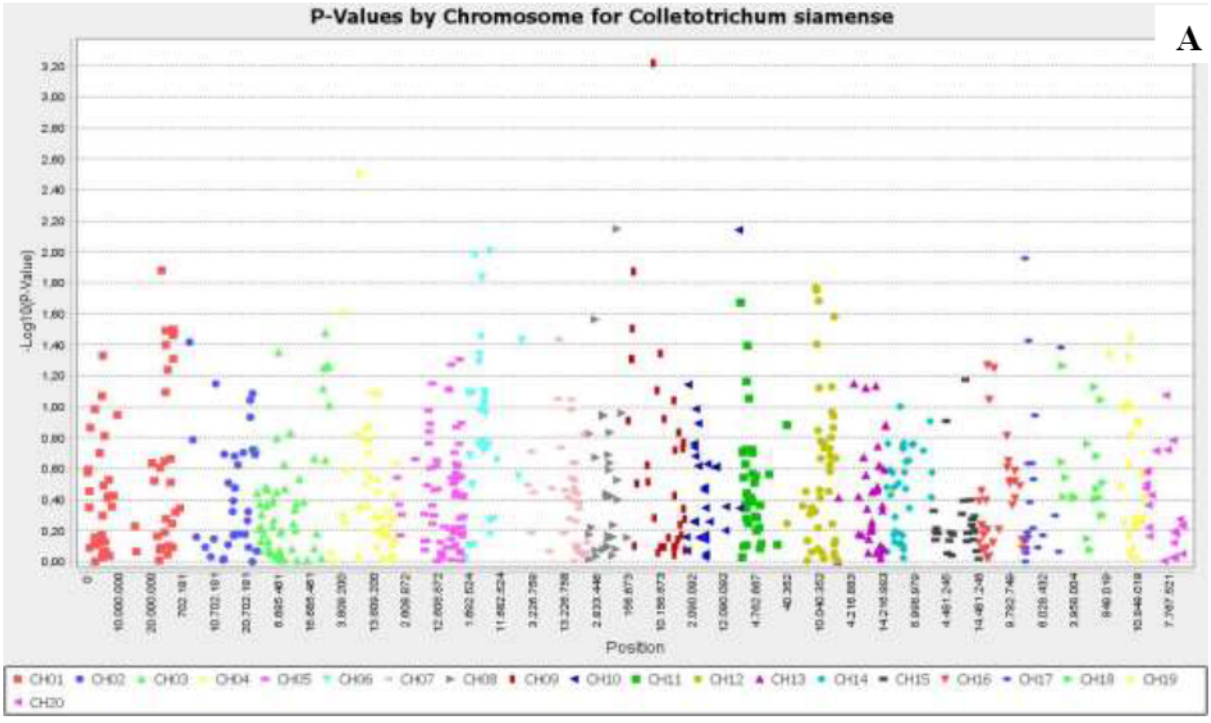
SNP loci with significant associations (P-value < 0.0097) and coefficient of determination (R^2) by the General Linear Model (GLM) and Mixed Linear Model (MLM) for Colletotrichum siamense resistance in the 'Haden' × 'Tommy Atkins' pseudo-F_2 population.*

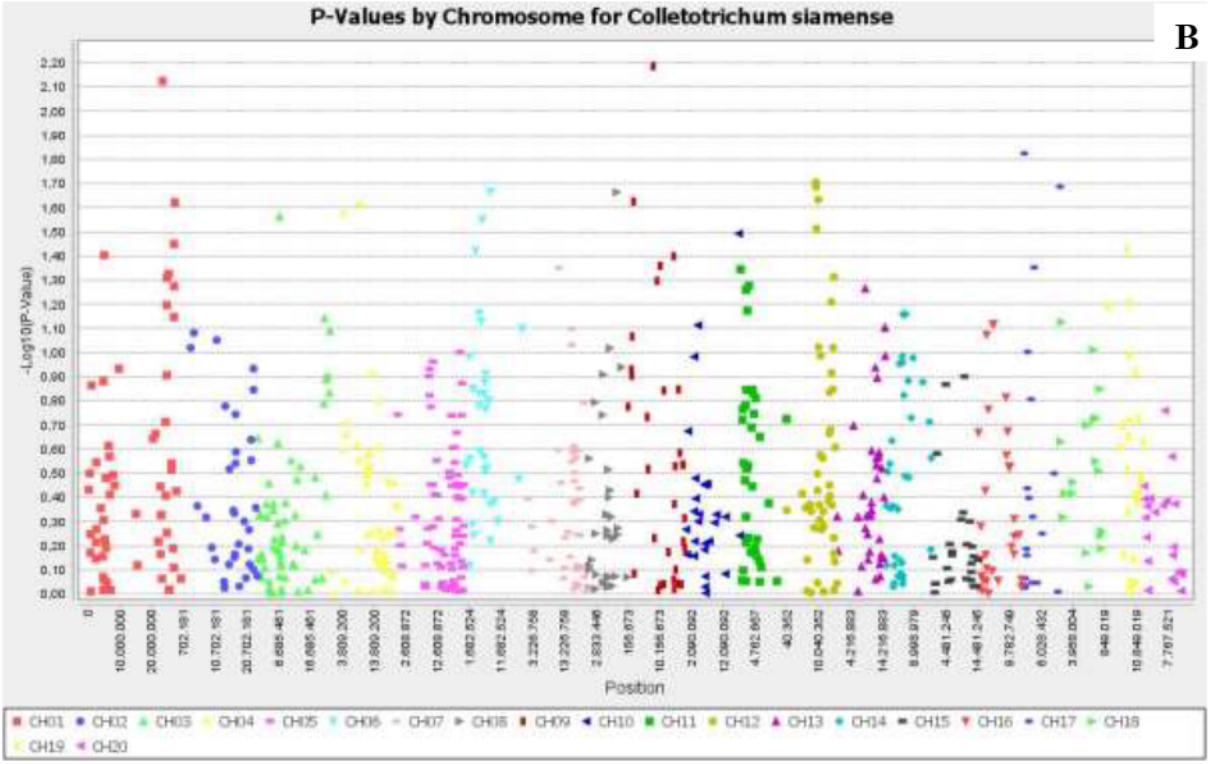
SNP	Chromosome	Position (bp)	*P-value*	Permutation[#]	R^2
GLM – Without Principal Components stratification					
Mi_0187	9	8122380	0.000605	0.26	14.24
Mi_0254	4	9321622	0.003131	0.78	10.52
Mi_0278	8	9348841	0.007117	0.96	8.81
Contig5471	10	16993797	0.007226	0.96	12.63
GLM – With Principal Components stratification					
Mi_0253	1	23533399	0.004825	0.41	9.00
Mi_0185	17	3928212	0.007537	0.57	11.00
Mi_0210	17	12070428	0.008098	0.60	8.00
SSKP030C1_G486T	11	2808983	0.008800	0.64	7.93
Mi_0187	9	8122380	0.009057	0.65	8.00
MLM – Without Principal Components stratification					
Mi_0187	9	8122380	0.006507	-	9.83
Mi_0253	1	23533399	0.007512	-	9.41
MLM – With Principal Components stratification					
Mi_0253	1	23533399	0.004441	-	10.64
Mi_0210	17	12070428	0.008789	-	8.96

*Percentage data for symptoms of *Colletotrichum siamense* transformed to arcsine
[#]Probability to a 1000 permutations.

Figure 1

Manhattan plots for SNPs associated with Colletotrichum siamense resistance in a 'Haden' × 'Tommy Atkins' mango pseudo-F_2 population: (A) General (GLM), (B) Mixed Linear Model (MLM).





According to Grinde *et al*. (2024), adjusting for principal component stratification (PCS) can induce spurious associations in genome-wide association studies (GWAS) of admixed populations, e.g., the 'Haden' × 'Tommy Atkins' cross in this study. In our analysis, two SNP loci were consistently identified in the GLM without PCS compared to allelic and genotypic tests (Table 1 and 2), indicating that GWAS performed without PCS may be more reliable.

Loci Mi_0187, Mi_0254, Mi_0278, and Contig5471, distributed across four chromosomes, explained between 8.81% and 14.24% of the total variance in GLM analysis without PCS, with an average of 11.55% and a combined variance of 46.2% for anthracnose resistance in the mango population (Table 2). The total variance reported here exceeds the 36.2% to 28.7% variance reported by Santos *et al*. (2023) in GWAS for dieback fungi in the same progenies.

Loci Mi_0187 and Mi_0254, on chromosomes 9 and 4, respectively, were significantly associated with resistance to *C. siamense* in both the Cochran-Armitage trend genotypic test (resistance >75%) and GLM and MLM analyses, except in MLM with PCS (Table 1). These loci, which control up to 24.76% of the total variance (Table 2), are strong candidates for further investigation and potential application in marker-assisted selection (MAS) for mango. Additionally, the Mi_0092 and Mi_0169 loci, identified only in allelic and genotypic analyses and located in close proximity on chromosome 16, are also candidates for further study, as they control approximately 7.11% of the variance (*P-value* = 0.056). Given their proximity, it is

reasonable to assume that both loci contribute 7.11% to resistance, likely representing a key chromosomal region for anthracnose resistance.

Aquino *et al*. (2021) reported high broad-sense heritability (~90%), estimated using REML/BLUP, in the same pseudo-F2 population, suggesting that anthracnose resistance in mango is controlled by a small number of genes, consistent with the few SNP loci identified in our study.

Along with mangosteen and guava, mango ranked as the 6th most produced fruit crop globally, with 59.15 million tons produced in 2022 (FAOSTAT, 2024). Thousands of mango cultivars exist worldwide, but only a few are traded internationally (Kuhn *et al*., 2017), primarily 'Tommy Atkins' and 'Palmer,' selected in Florida, USA. The limited number of internationally traded cultivars is due to long juvenile periods, large plant sizes, small breeding populations, and limited use of molecular markers for MAS. However, the availability of numerous SNP markers (Kuhn *et al*., 2016; Srivastav *et al*., 2023) and the mango genome sequence (Wang *et al*., 2020) enables GWAS and QTL analyses, with the ultimate goal of applying MAS to accelerate the development of new cultivars.

According to a review by Song *et al*. (2023), studies on mango gene functions have focused on genes related to flowering regulation, fruit development, and polyphenol biosynthesis. However, research on genes associated with major mango diseases has been limited. To our knowledge, this is a pioneering GWAS applied to anthracnose, the most important and widespread fungal disease affecting mango fruit worldwide, including commercial orchards in Northeast Brazil, a leading region for mango production (Lima *et al*. 2013).

As emphasized by Santos *et al*. (2023), coding for 1 (controls) and 2 (cases) in disease percentage data simplifies the complex plant-host interaction. However, this approach enables the identification of consensus chromosomal regions associated with anthracnose, as demonstrated by the significant association of SNP loci through both qualitative (Fisher's exact test and Cochran-Armitage trend genotypic test) and quantitative (GLM and MLM) analyses.

Loci Mi_0187 and Mi_0254, identified in both qualitative and quantitative analyses, are strong candidates for designing PCR primers to implement MAS in mango, aiming to select genotypes with enhanced resistance to anthracnose. Additionally, the Mi_0092, Mi_0169, Mi_0278, and Contig5471 loci, identified in qualitative and quantitative analyses, are also potential candidates, though with slightly lower potential, for further analysis to apply MAS in mango. Collectively, these six SNP loci account for up to 53.31% (R^2) of the total variance in anthracnose resistance in mango.

5 CONCLUSION

The combination of qualitative and quantitative genome association methods enabled the identification of two SNP loci that control up to 24.76% of the total variance in anthracnose resistance in mango. An additional four SNP loci, identified exclusively through either qualitative or quantitative analyses, explain 28.55% of the total variance. These SNP loci are strong candidates for the design of PCR primers, facilitating the application of marker-assisted selection in mango breeding programs.

ACKNOWLEDGMENTS

We extend our thanks to David Kuhn and Barbie Freeman (USDA-ARS-SHRS, USA) for providing the genotyping data for the 'Haden' × 'Tommy Atkins' pseudo-F2 population. CAF Santos is a CNPq fellowship researcher.

REFERENCES

Anvisa. *Agência Nacional de Vigilância Sanitária. Resolução-RDC Nº 61, de 3 de Fevereiro de 2016*. (2016) Available at <https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/22301726/do1-2016-02-05-resolucao-rdc-n-61-de-3-de-fevereiro-de-2016-22301384>.

Aquino, D.A.L. Santos, C.A.F., Batista, D.C., Camara, M.P.R. (2021).Progeny selection and inheritance of resistance to *Colletotrichum siamense* in *Mangifera indica* crosses. *Fruits*, 76:230-235.

Aruz, L.F. (2000). Mango anthracnose: economic impact and current options for integrated and management. *Plant Disease*, 84(6):600-611.

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., Buckler, E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23:2633-2635.

Brasil. 2022. Ministério da Agricultura, Pecuaria e Abastecimento (MAPA). https:// indicadores.agricultura.gov.br/agrostat/index.htm (accessed 10 October 2024).

Dautt-Castro, M., López-Virgen, A.G., Ochoa-Leyva, A., Contreras-Vergara, C.A., Sortillón-Sortillón, A.P., Martínez-Téllez, M.A,, González-Aguilar, G.A., Casas-Flores, J.S., Sañudo-Barajas, A., Kuhn, D.N. (2019). Genome-Wide Identification of Mango (*Mangifera indica* L.) Polygalacturonases: Expression Analysis of Family Members and Total Enzyme Activity During Fruit Ripening. *Frontier of Plant Science*, 10:969.

Duval, M.F., Bunel, J., Sitbon, C., Risterucci, A.M. (2005). Development of microsatellite markers for mango (*Mangifera indica* L.). *Molecular Ecology Notes* 5:824–826.

FAOSTAT (2024). *Food and agriculture Organization United Nations.* Available at https://www.fao.org/faostat/en/#data/TCL/visualize.

Grinde, K.E., Browning, B.L., Reiner, A.P., Thornton, T.A., Browning, S.R. (2024) Adjusting for principal components can induce spurious associations in genome-wide association studies in admixed populations. *bioRxiv [Preprint*]. 2024 Apr. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11014513/

Gupta, S., Singh, K. P., & Singh, A. K. (2015). Resistance to anthracnose disease in commercial cultivars to hybrids of mango. *Plant Pathology Journal,* 14: 225-258.

Korte, A., Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9:29.

Kuhn, D.N., Bally, I.S.E., Dillon, N.L., Innes, D., Groh, A.M., Rahaman, J., Ophir, R., Cohen, Y., Sherman, A. (2017). Genetic Map of Mango: A Tool for Mango Breeding. *Frontier of Plant Science*. 8:577.

Kuhn, D.N., Dillon, N., Innes, D., Wu, L-S., Mockaitis, K. (2016). Development of single nucleotide polymorphism (SNP) markers from the mango (*Mangifera indica*) transcriptome for mapping and estimation of genetic diversity. *Acta Horticulture*, 1111:315–322.

Lima, N.B., Batista, M.V.A., Morais, M.A., Barbosa, M.A.G., Michereff, S.J., Hyde, K.D., Câmara, M.P.S. (2013). Five *Colletotrichum* species are responsible for mango anthracnose in northeastern Brazil. *Fungal Diversity,* 61:75-88.

Lima, N.B., Lima, W.G., Tovar-Pedraza, J.M., Michereff, S.J., Câmara, M.P. (2015) Comparative epidemiology of *Colletotrichum* species from mango in northeastern Brazil. *European Journal of Plant Pathology*, 141(4):679-688.

Ma, X., Wu, H., Liu. B., Wang, S., Zhang, Y., Su, M., Zheng, B., Pan, H., Du, B., Wang, J., He, P., Chen, Q., An, H., Xu, W., Luo, X. (2024). Genomic diversity, population structure, and genome-wide association reveal genetic differentiation and trait improvements in mango. *Horticulture Research*, 11(7):uhae153.

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature* 456, 18–21.

Majumder, P.P., Ghosh, S. (2005). Mapping quantitative trait loci in humans: achievements and limitations. *Journal of Clinical Investigation*, 115(6):1419-24.

Marees, A.T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., Derks, EM. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal Methods Psychiatric Research*, 27(2):e1608.

Ogata, T., Yamanaka, S., Shoda, M., Urasaki, N., & Yamamoto, T. (2016). Current status of tropical fruit breeding and genetics for three tropical fruit species cultivated in Japan: pineapple, mango, and papaya. *Breeding Science,* 66:69-81.

Ploetz, R.C. (2010). *Anthracnose of mango: Management of the most important pre-and post-harvest disease.* Report: University of Florida, TREC-Homestead. Departament of Plant Pathology, p. 1-11.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., Bakker, P.I.D., Daly, M.J. *et al*. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetic*, 81:559–575.

Santos, C.A.F.; Aquino, D.A.L.; Batista, D.C.; Alves, E.O.S. (2023). Genomic association of resistance to dieback diseases in a pseudo-F2 mango population. *Genetics and molecular research*, 22: 1-11.

Song, M., Wang, H., Fan, Z., Huang, H., Ma, H. (2023). Advances in sequencing and key character analysis of mango (*Mangifera indica* L.). *Horticulture Research*, 10(2):1-16.

Srivastav, M., Radadiya, N., Ramachandra, S., Jayaswal, P.K., Singh, N., Singh, S., Mahato, A.K., Tandon, G., Gupta, A., Devi, R., Subrayagowda, S.H., Kumar, G., Prakash, P., Singh, S., Sharma, N., Nagaraja, A., Kar, A., Rudra, S.G., Sethi, S., Jaiswal, S., Iquebal, M.A., Singh, R., Singh, S.K., Singh, N.K. (2023) High resolution mapping of QTLs for fruit color and firmness in Amrapali/Sensation mango hybrids. *Frontier of Plant Science*. 7:14-1135285.

Terao, D., Batista, D.C., Ribeiro, I.J.A. (2016). Doenças da mangueira (*Mangifera indica.* L). In H. Kimati, L. Amorin, J. A. M. Rezende, A. Bergamin Filho & L. E. A. Camargo (5ª eds.), *Manual de fitopatologia: doenças das plantas cultivadas*. São Paulo, Brasil. pp. 224-533.

Tucho, A., Lemessa, F., Berecha, G. (2014). Distribution and Occurrence of Mango Anthracnose (Colletotrichum gloesporioides Penz and Sacc) in Humid Agro-ecology of Southwest Ethiopia. *Plant Pathology Journal,* 13:268-277.

Uffelmann, E., Huang, Q.Q., Munung, N.S. *et al*. (2021). Genome-wide association studies. *Nature Review Methods Primers* 1, 59.

Wang, P; Luo, Y; Huang, J; Gao, S; Zhu, G; Dang, Z; Gai, J; Yang, M; Zhu, M; Zhang, H. (2020) The genome evolution and domestication of tropical fruit mango. *Genome Biology*, 21:60p.