

A SNP-Based Linkage Map and QTL Identification for Resistance to Yam Anthracnose Disease (YAD) in Water Yam (*Dioscorea alata*)

Kwabena Darkwa

International Institute of Tropical Agriculture

Paterne AGRE (✉ P.Agre@cgiar.org)

International Institute of Tropical Agriculture, IITA, Ibadan, Nigeria <https://orcid.org/0000-0003-1231-2530>

Bunmi Olasanmi

Department of Agronomy, University of Ibadan, Nigeria

Olufisayo Kolade

International Institute of Tropical Agriculture

Pierre Moumet

CIRAD, UMR, AGAP, F-34398 Montpellier, France

Antonio Lopez-Montes

International Trade Center, Accra, Ghana

David Dekoeyer

Agriculture and Agri-Food, Canada, Fredericton

Patrick Adebola

International Institute of Tropical Agriculture, Abuja Nigeria

Lava P. Kumar

International Institute of Tropical Agriculture

Robert Asiedu

International Institute of Tropical Agriculture

Asrat Asfaw

International Institute of Tropical Agriculture

Research article

Keywords: *Dioscorea* spp, greater yam, genetic map, marker-trait association, linkage analysis

Posted Date: May 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-26806/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Yam anthracnose disease (YAD) caused by *Colletotrichum gloeosporioides* is the primary cause of yield loss in water yam (*Dioscorea alata*), the widely cultivated species of yam. Development of resistant cultivars have been a prime target for sustainable management of anthracnose in water yam. Molecular breeding tools are required to expedite the development of improved yam varieties. QTL analysis using high density genetic maps serve as a powerful tool to discover key locations of quantitative traits. This study aimed at tagging quantitative trait loci (QTL) for anthracnose disease resistance in a biparental mapping population of *D. alata*.

Results: In this study, two contrasting parents for yam anthracnose disease reaction and their 204 full-sib offspring were used to develop a high-density genetic linkage map with 3,257 SNP markers by the GBS technique. The total length of the consensus map was 1460.94 cM with an average of 163 markers per chromosome. Four QTLs were detected for anthracnose disease resistance in 4 locations on 3 chromosomes. The proportion of phenotypic variance explained by these QTLs ranged from 10 to 13%. Plant defense response genes including GDSL-like Lipase/Acylhydrolase, Protein kinase domain and F-box protein were also detected within the QTL regions.

Conclusion: The results from the present study provide valuable insight into the genetic architecture of anthracnose resistance in water yam. The candidate markers and putative genes identified herewith form a relevant resource to apply marker-assisted selection as alternative to a conventional labor-intensive screening for anthracnose resistance in water yam.

Background

Yam (*Dioscorea* spp.) is a multi-species monocotyledonous crop widely grown in the tropics and subtropics [1]. It is the most valuable crop in West Africa, where its cultivation began 11,000 years ago [2]. Of the over 600 yam species, water yam (*D. alata*) is the extensively cultivated species worldwide [3]. In Africa, white yam (*D. rotundata*) is the most cultivated yam species followed by water yam [3]. In West Africa, yam production is mainly by smallholder farmers, making it a major source of farm employment and income for this group. In addition, yam plays significant role in traditional medicine and in the socio-cultural life of the people as it is involved in many key life ceremonies [4].

Water yam possess a number of valuable attributes for cultivation and consumption. These include high multiplication ratio, early vigour for weed smothering, higher genetic potential for high yield (especially under low to average soil fertility), low post-harvest losses, good processing quality and high nutritive value including its possession of low glycemic index [5; 6]. However, anthracnose disease caused by the *Colletotrichum gloeosporioides* (Penz) is the most limiting factor affecting productivity of water yam in many regions of the world [7]. Anthracnose causes mild to acute leaf necrosis, premature leaf abscission and shoot die-back [8]. Severe infections result in defoliation leaving naked, black and drying vines [9]. Yield losses from the disease of up to 90% have been reported under severe conditions on different cultivars of water yam in West Africa, Central America and the Pacific [10, 11, 12, 13]. High genetic and

pathogenic variance has been reported among isolates of *C. gloeosporioides* from different geographical locations [7, 14, 15], suggesting that there is high probability for the geographic variation in strains, some of which could overcome existing resistance [16].

Cultural control approaches such as the use of disease-free planting materials, adjustment of plant spacing and planting dates, burying infected plant residues in the soil immediately after harvesting, intercropping, crop rotation with non-host crops and fallowing have been used in other plant pathosystems to reduce pathogen inoculum in the field, delay disease onset, or slow disease progress [17, 18]. Nonetheless, these disease management practices have not been effective for controlling anthracnose disease in water yam or result in substantial increase in tuber yield [19], especially in disease endemic areas. Also, biological control to impede or outcompete the multiplication and spread of virulent *C. gloeosporioides* strains in yam fields has been limited [20]. Chemical control can be an effective disease management approach but most yam producers are smallholder growers and may not have the prerequisite technical support and finance to afford the use of fungicides [21]. Furthermore, inappropriate use of fungicides could potentially result in the development of resistant *C. gloeosporioides* strains to systemic fungicides [22] as well as detrimental environmental effects. The best control option is therefore the development and deployment of anthracnose resistant water yam varieties. It is, therefore, expedient to develop varieties with multiple disease resistance genes to provide stable and durable resistance against the broad spectrum of the fungal pathogen.

Substantial progress has been made to develop anthracnose resistant water yam varieties at the International Institute of Tropical Agriculture (IITA), Nigeria and national agricultural research systems in West Africa and elsewhere through conventional breeding using phenotypic observations. This effort is, however, arduous and considerably slow due to the inherent biological constraints of a heterozygous vegetatively propagated crop [23]. Genomics-informed breeding techniques such as molecular marker assisted breeding and genomic selection would accelerate efforts in introgressing anthracnose resistance into preferred genetic backgrounds [3].

Earlier investigations on anthracnose disease in water yam showed that resistance is likely to be dominant and quantitatively inherited [24]. Efforts have also been made to identify QTL controlling YAD using low-throughput molecular markers and less dense or unsaturated genetic maps such as AFLP markers [5, 25] and EST-SSRs [26]. Prospects for locating additional QTL and applying molecular breeding methods in water yam improvement programs is very promising especially due to advances in next-generation sequencing and the recent development of the reference genome sequence of *D. rotundata* and *D. alata*. The objective of this study was to develop a SNP-based genetic linkage map and identify QTL for anthracnose disease resistance in a bi-parental mapping population of *D. alata*.

Results

Phenotypic variability of the parents and derived clones

Significant differences ($p < 0.05$) were observed among the mapping population for their reaction to YAD in both years (Table 1). The mean squares for year as well as genotype by year interaction were highly significant ($p < 0.01$). The disease pressure was higher in 2018 compared to 2017. The area under disease progression curve (AUDPC) estimates ranged from 210.0 to 397.5 with an average of 245.5 in 2017 while the range was from 233.4 to 482.1 with an average of 299.8 in 2018. Exposure of the progenies to the natural field infestation by anthracnose revealed none of the lines were highly resistant (mean severity score of 1, equivalent to AUDPC value < 105) or highly susceptible (mean severity score of 5, equivalent of AUDPC > 525) (Fig. 1). Majority of the genotypes evaluated (67–92%) were moderately resistant to anthracnose across the 2 years. Broad sense heritability for YAD resistance was high (70.64%).

Table 1

Mean squares and heritability estimate for yam anthracnose disease severity in the mapping population evaluated in 2017 and 2018 at Ibadan, Nigeria

Trait	Mean Squares			CV	Broad sense heritability
	Genotype	Year	Genotype x Year	(%)	(%)
AUDPC	163.01 *	2190.01**	3371.8***	17.6	70.64

AUDPC area under disease progression curve, ** significant at 1%, * significant at 5% probability level, CV coefficient of variation

Genetic Map

Chi-square test with an initial set of 15,936 SNP markers identified 7,583 SNPs (47.6%) as informative markers for linkage map construction. Further filtering for 20% missing data (both for the SNP and Genotype) (Fig. 2) followed by conservative Bonferroni threshold on the informative SNP markers identified 3,931 markers as distorted and 395 as redundant and were therefore removed before the map construction (Fig. 3). Prior the genetic map construction, Pairwise recombination fractions for all markers were calculated and SNP marker were ordered (Supplementary file Fig. 1). The final genetic map was constructed using the highly informative 3,257 SNP markers that covered all 20 linkage groups of the water yam genome (Fig. 4, Supplementary Fig. 2), with a total genetic distance of 1460.94 cM. This linkage map had an average of 163 markers per linkage group or chromosome. Linkage group 6, 5 and 2 were the longest with 109.52 cM, 109.19 cM and 109.17 cM, respectively, while linkage group 19 was the shortest with 33.08 cM (Table 2). Genetic gap across the genome varied from 2.39 cM on linkage group 15 to 44.44 cM on linkage group 2 which reflected map uniformity (Table 2).

Table 2
Distribution of SNP markers and linkage group length in each chromosome

Chromosomes	Number of SNPs	Chromosome Length (cM)	Maximum Gap
Chr1	80	80.92	21.63
Chr2	84	109.17	44.44
Chr3	115	64.62	12.66
Chr4	548	92.32	8.26
Chr5	219	109.19	15.95
Chr6	191	109.52	16.22
Chr7	127	59.84	14.54
Chr8	206	91.38	11.56
Chr9	124	77.12	16.75
Chr10	104	57.13	6.52
Chr11	85	55.95	9.78
Chr12	125	83.21	30.84
Chr13	116	95.55	15.63
Chr14	208	51.46	6.64
Chr15	189	34.70	2.39
Chr16	139	77.16	10.82
Chr17	208	67.44	13.07
Chr18	129	75.70	15.82
Chr19	111	33.08	5.66
Chr20	149	35.48	5.89
Total	3,257	1460.94	

Qtl Analysis With The Genetic Map

The QTL detected on three chromosomes out of the 20 for the specific year, as well as across years data are presented in Table 3 and Fig. 5. One significant QTL was detected on chromosome 7 at position 10.60 cM (between 10.596 and 19.217 cM) for the 2017 data. This QTL (chr7_3179921) had LOD score of 4.51 and accounted for 11.83% of the total phenotypic variation in anthracnose severity score (Table 3). Another significant QTL was detected on chromosome 15 with 2018 data. The QTL (Chr15_8632438) was

at position 28.80 cM (between 10.171 and 28.817 cM) at LOD score of 4.43 and contributed 10.73% of the total phenotypic variance.

QTL analysis using the averaged data on AUDPC estimates across the two years detected two additional QTL for this trait on chromosomes 7 and 18. The QTL on chromosome 7 (chr07_5765300) was at position 19.21 cM (between 10.596 and 19.218 cM) at LOD score of 5.28 and accounted for 10.04% of the total phenotypic variation in anthracnose severity score. The QTL on chromosome 18 (chr18_18405143) was at position 61.4 cM (between 61.345 and 61.432 cM) at LOD score of 4.65 and contributed 12.59% of the total phenotypic variance.

Table 3
Summary of significant QTLs detected for yam anthracnose disease resistance in water yam.

Year	Markers	Chr	Pos (cM)	LOD	P value	Cl. Low	Cl. High	R ² (%)
2017	chr7_3179921	7	10.60	4.51	5.434e-05	10.596	19.217	11.83
2018	chr15_8632438	15	28.80	4.43	7.17e-03	10.171	28.817	10.73
Average	chr07_5765300	7	19.21	5.28	2.60e-4	10.596	19.218	10.04
	chr18_18405143	18	61.4	4.65	2.769e-05	61.345	61.432	12.59
Chr = chromosome, pos = position, LOD = logarithm of odds score, CI = confidence interval, R ² = % variation explained.								

Gene annotation within the QTL regions of the significant QTLs identified putative plant defense response genes. Adjacent chr7_3179921 were two plant biotic stress related genes; DRNTG_08663.1 (GDSL-like Lipase/Acylhydrolase) and DRNTG_08664.1 (Protein kinase domain). The N-terminal alpha/beta domain gene (DRNTG_14305.1) was detected within the flanking sequence of chr15_8632438. Two additional genes DRNTG_18245.1 (ANTH domain - Putative clathrin assembly protein) and DRNTG_29617.1 (WD domain - WD40 repeat-containing protein) were detected within the vicinity of chr18_18405143, while the F-box protein (DRNTG_23336.1) was found near chr07_5765300.

Interaction among the four QTL detected for yam anthracnose disease resistance revealed significant ($p < 0.05$) QTL by QTL interaction for chr7_3179921 and chr15_8632438, and chr07_5765300 and chr18_18405143, while no significant variation was observed among all other QTL combinations (Table 4).

Table 4
Interactions among the detected QTL

Trait	Markers interaction	df	MS	P-value	Adjusted R-squared
AUDPC	Chr7@10.60:Chr7@19.21	1	55.9	0.835	0.04147
	Chr7@10.60:Chr15@28.8	1	5303.5	0.0456*	0.02544
	Chr7@10.60:Chr18@61.4	1	155.2	0.734	-0.0002131
	Chr7@19.21:Chr15@28.8	1	2580.7	0.158	0.04395
	Chr7@19.21:Chr18@61.4	1	6341.0	0.026*	0.06074
	Chr15@28.80:Chr18@61.4	1	1408.4	0.309	-0.01079
	Chr7@10.60:Chr7@28.8:chr15@28.8:chr18@61.4	3	1247.7	0.068	0.04413
	df: degree of freedom, MS: mean square				

Marker Effect

The presence of allele T for marker chr7_3179921 and chr7_5765300 appears to lower the AUDPC in the population evaluated in this study, while the presence of the C allele tends to increase the AUDPC value for yam anthracnose disease (Figs. 6 and 7). Presence of allele A on the variant AG/GA tends to increase the AUDPC value while the presence of G lowers the AUDPC value in the population (Figs. 8 and 9).

Discussions

Anthrachnose is one of the major constraints contributing to yield reduction in water yam production. It occurs wherever water yam grows. Several management options are available to tackle the anthracnose disease threat in water yam production. However, development of cultivars with stable and durable resistance against the broad spectrum of the fungal pathogen is the best option for integrated management of the disease. This study assessed functional association of anthracnose resistance and genetic markers via QTL mapping approach using recombinant clonal population to enable indirect selection for resistance to the disease in cultivar development efforts. The recombinant clonal population showed differential response to the disease-causing organism across the two-year evaluation period. The population showed quantitative resistance with a continuous distribution from resistance to susceptible range with substantial skewness towards resistance. No immune lines were identified through the natural field exposure to *C. gloeosporioides*, rather, large number of the lines were resistant or moderately resistant. Result of this study confirms the earlier reports by Mignouna et al. [5], Petro et al. [25] and Bhattacharjee et al. [26], that resistance to YAD is dominantly and qualitatively inherited trait. The heritability estimate in the present study was high just as earlier reported by Petro et al. [25] and Bhattacharjee et al. [26].

The high-density genetic linkage map using 3,257 SNPs from the GBS platform that spanned a total length of 1460.94 cM represents the most saturated genetic map for *D. alata* to date. In an earlier effort, Cormier et al. [27] constructed a high-density genetic map of *D. alata* using 1,579 polymorphic SNP markers with a consensus map length of 2613.5 cM leading to the identification of a major QTL for sex determination on linkage group six. Genetic linkage maps of water yam were also developed using EST-SSRs [26] and AFLPs [5; 25]. The genetic linkage map presented in this report will offer an outstanding genetic background for qualitative and quantitative trait analysis of water yam.

Three studies have so far been conducted for mapping QTLs controlling resistance to anthracnose in water yam [5; 25, 26]. The study by Mignouna et al. [5] and Petro et al. [25] utilized AFLP maps and identified one and nine QTLs, respectively for anthracnose resistance explaining 10% and 26 to 74% of the total phenotypic variation. Bhattacharjee et al. [26] utilized an EST-SSR genetic map for their study and identified a major QTL on linkage group 14 explaining 69% of the total phenotypic variance. Even though the previous studies ordered markers on 20 linkage groups, the absence of a common genetic map and the different marker systems makes it difficult to compare the location of the detected QTLs in these studies. In this study, four QTLs for yam anthracnose disease resistance were identified on four chromosomes explaining 10.0 to 12.6% of the total phenotypic variation in the trait. The detected QTL showed general or year specific effect which could be attributed to variation in strain or intensity of strains of *C. gloeosporioides* infestation over years during the field experimentation. Earlier investigations have also demonstrated resistance of water yam to anthracnose disease to be both isolate-specific and non-specific [24; 25]. Strain-specific as well as non-strain-specific QTLs were detected for anthracnose resistance in different populations of *D. alata* [5; 25]. Also, Geffroy et al. [28] reported QTL specific for anthracnose resistance in leaves, stems and petioles of common bean. A similar mechanism of isolate specificity or non-specificity may have occurred in this study leading to the identification of different QTL conferring anthracnose resistance over the two years.

Gene annotation identified plant biotic stress response genes within the flanks of the significant QTL for YAD discovered in this study. The GDSL-like Lipase/Acylhydrolase gene detected within the vicinity of chr7_3179921 was reported to regulate systemic resistance to *Alternaria brassicicola* in Arabidopsis [29; 30]. Hong et al. [31] also found this gene to be involved in the defense against drought and *Xanthomonas campestris* pv. Vesicatoria in pepper. The protein kinase domain implicated in the resistance against bacterial blight (*Xanthomonas oryzae*) in rice [32] and resistance to the necrotrophic fungal pathogen, *Plectosphaerella cucumerina* in Arabidopsis [33] was also found within the QTL region of chr7_3179921.

The ANTH domain and WD domain discovered within the vicinity of chr18_18405143 were reported to be important in *Nicotiana benthamiana* and Arabidopsis defense against *Pseudomonas syringae* [34] and enhanced the resistance to anthracnose leaf blights in maize caused by *Colletotrichum sublineolum* [35; 36], respectively. The F-box protein found within the QTL region of chr07_5765300 was reported to be involved in cell death and defense response during the pathogen recognition of *Pseudomonas syringae* and Tobacco mosaic virus in tomato and tobacco [37]. The N – terminal domain identified within the flanking sequence of chr15_8632438 in the present study was reported to be involved in the resistance of

Arabidopsis to the downy mildew pathogen *Hyaloperonospora arabidopsidis* [38]. Therefore, enough evidence exists that the genes within the flanks of the significant QTL for anthracnose disease resistance discovered in this study code for response to plant biotic stress.

Conclusion

In this study, we report the use of a highly saturated genetic linkage map based on SNP markers to identify QTLs for YAD in water yam. The high-density genetic map based on 3,257 SNPs would be very useful for genetic studies in water yam. The linkage analysis identified four QTLs accounting for 10.0 to 12.6% of the total phenotypic variance in anthracnose severity score. Five genes involved in plant defense against diseases were also identified within the flanks of the QTLs detected in this study. These results, upon validation and development of diagnostic SNP markers will enable the application of marker-assisted selection especially at the early breeding stages to shorten the breeding cycle of water yam. Giving the quantitative nature of this trait and the proportion of the phenotypic variance unaccounted for in this studies, future investigations should consider larger mapping populations through meta-QTL analysis encompassing diverse genetic backgrounds with high level of resistance and high through-put phenotyping for precise disease evaluation across different locations and years.

Materials And Methods

Plant materials

A recombinant clonal population of 204 genotypes from the cross of TDa9900015 x TDa0500048 developed at IITA was used for this study. TDa9900015 is a female breeding line moderately resistant to YAD while TDa0500048 is a male breeding line expressing susceptible reaction to the disease. All the recombinant clones plus both parents and one highly susceptible check (variety# TDa92-2) were evaluated in field experiment. The field anthracnose phenotyping experiment was conducted in two seasons (2017 and 2018) at IITA, Ibadan research farm in Nigeria. The field experiment was carried out using partial replicated design during the major rainy seasons when anthracnose incidence and severity are high. The susceptible check (TDa92-2) was planted as spreader row between blocks and around the field.

Phenotyping

Anthracnose disease severity was scored at two months after planting and thereafter, fortnightly till six months. Severity was scored by visual assessment of the relative area of plant tissue affected by anthracnose using a 1–5 severity rating scale. Where, 1 = No visible symptoms of anthracnose disease or infection spot on the leaf surface; 2 = Few anthracnose spots or symptoms on 1 to 25% of the plant (i.e. one or two spots of less than 1 cm diameter width, and dry tissue on the leaf surface); 3 = Anthracnose symptoms covering 26 to 50% of the plant (i.e. one or two spots of more than 1 cm diameter width, and dry tissue on the leaf surface, small dark and no dried spots with more than 1 cm width are present); 4 = Symptoms on > 50% of the plant (i.e. coalesced spots with dry tissue and covering a significant

proportion of the leaf surface, areas with less than 1 cm width coalesce to bigger spots and yellowing of green tissue is intense around the spots areas); and 5 = Severe necrosis and death of the plant (i.e. coalesced spots with dry tissue more than 1.5 cm in diameter and covering a great proportion of the leaf surface and yellowing of the green tissue is generalized in the leaf blade) [39].

The area under the disease progression curve (AUDPC) was estimated from the disease severity scores using the trapezoidal method [40]. This method discretizes the time variable and calculates the average disease intensity between each pair of adjacent time points. (see Equation 1 in the Supplementary Files)

$$AUDPC = \sum_{i=1}^n \left(\frac{y_i + y_{i+1}}{2} \right) (t_{i+1} - t_i)$$

where n = total number of observations, y_i = disease severity at the i th observation, and t = time at the i th observation.

Genotyping

Young fresh leaf samples were collected from the 206 genotypes (204 recombinant progenies, the two parents and a check variety) and immediately dipped in dry ice. The leaves were stored at -80°C prior to lyophilization. Lyophilized leaf samples were sent to CIRAD-France for DNA extraction, library construction and Genotyping by Sequencing (GBS). DNA extraction and Genotyping by Sequencing (GBS) were performed as described in Cormier et al. [29]. GBS libraries were constructed as described by Elshire et al. [41] using PstI-MseI restriction enzymes. Sequencing was conducted on Illumina HiSeq 3000 system (150 bp, single-end reads) at the GeT-PlaGe platform in Toulouse, France.

Data Analyses

Phenotype data

Anthracoze severity score data collected at different time during the crop's growth period were converted to AUDPC for quantitative comparison across years. The relative area under disease progress curve data was subjected to mixed model analysis using lme4 library package implemented in R (see Equation 2 in the Supplementary Files)[42].

$$Y_{ijk} = \mu + \beta_i + R_{ij} + G_k + (\beta_i \times G_k) + e_{ijkm}$$

Where; Y_{ijk} = phenotypic value, μ = overall phenotypic mean, β_i = effect of year i , R_{ij} = effect of block j in year i , G_k = effect of genotype k , $(\beta_i \times G_k)$ = effect of interaction between year i and genotype k and e_{ijkm} = residual. Block effects were added to the model as random variable to remove the spatial variation within the trial field. Broad sense heritability was estimated from the model to assess the proportion of

phenotypic variation in data set due to genetic effects. Phenotypic BLUE (Best Linear Unbiased Estimator) values of un-shrunken means for QTL analysis were extracted for the years and across years.

SnP Calling And Quality Assessment

Raw data was first filtered using a pipeline described in Scarcelli et al. [43]. Demuladapt (<https://github.com/Maillol/demultadapt>) was used for demultiplexing and cutadapt 1.2.1 [44] used to remove the adaptors and low-quality bases reads with a mean quality score < 30 using a free perl script https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl. For the final SNP calling, GATK was used while mapping was performed using default options of Burrows-Wheeler Aligner (BWA) [45] using the *D. alata* reference genome. The SNP quality assessment was performed using vcftools [46] and plink [47], and SNPs with low information (such as MAF < 0.05, depth of sequencing < 5, 20% missing data for both SNPs and genotypes, multiple alleles) were removed.

Genetic Map Construction

The VCF file with the filtered SNPs including the two parents alongside the progenies was used for the linkage analysis using MAPpoly package [48] in R environment [42]. Chi-square test was applied for all the markers, considering the expected segregation patterns under Mendelian inheritance, random chromosome pairing and no double reduction using the filter.non.conforming function in MAPpoly library in order to remove non-informative SNP markers from further analysis. Linkage grouping was performed using an initial LOD value of > 4 in MAPpoly. The LOD value of 4.0 that established known linkage groups was then chosen as the significance criterion for multipoint linkage testing. Marker order and diagnostics were performed using the thresh.LOD.ph and thresh.LOD.rf function in MAPpoly R package. The final high-density genetic map was constructed using R/QTL2 [49] and viewed in LinkageMapView [50]. The Qtl/jittermap was used to adjust the genetic position of closely related makers.

Qtl Analysis And Annotation

The QTL analysis was performed using the Composite Interval Mapping (CIM) method in R/QTL2 package [49]. A forward and backward simple stepwise regression was run to select background markers having a significant level of $P < 0.10$. The threshold levels to declare significant QTLs were empirically determined through 1000 permutations of the data, which maintained chromosome-wise Type I error rate of 0.05 [51].

The location of a QTL was described according to its LOD peak location and the surrounding region with 95% confidence interval using the Bayesian model. The proportion of phenotypic variance accounted for by each detected QTL was estimated by a single-factor analysis of variance using General Linear Model

procedure on the individual marker loci closest to the QTL identified by CIM. Interaction among significant QTL was assessed using the R/QTL2 package [49]. The marker effect of each significant marker was estimated through a simple step wise regression analysis using lme4 R package. The gene-finding format files (gff) of the *D. rotundata* and *D. alata* reference genomes (https://drive.google.com/drive/folders/1H5T4xjKAEI9LiR-4qK_IR6TypCDe8nj and

https://yambase.org/organism/Dioscorea_alata/genome) were used to locate candidate genes within the flanking sequences (5Kb based-pair) of the QTL detected for anthracnose disease resistance.

Abbreviations

YAD Yam anthracnose disease

QTL Quantitative Trait Loci

SNPs Single nucleotide polymorphisms

AFLP amplified fragment length polymorphism

EST Expressed sequence tags

SSRs Simple Sequence Repeats

AUDPC Area Under Disease Progression Curve

LOD logarithm of odds score

Declarations

Acknowledgments

We acknowledge funding from IITA for the AfricaYam project through a research grant (OPP1052998) received from the Bill and Melinda Gates Foundation (BMGF). Special thanks to Ranjana Bhattacharjee for contributing to the population development. The African Union Commission is highly appreciated for funding Kwabena Darkwa at the Pan African University, Institute of Life and Earth Sciences (PAULESI). The yam research at IITA is supported by the CGIAR Research Program on Roots, Tubers and Bananas (RTB), supported by CGIAR Fund Donors.

Author contributions

Asrat Asfaw, Antonio Lopez-Montes, Robert Asiedu, Bunmi Olasanmi and David De koeyer designed the project; Asrat Asfaw and Antonio Lopez-Montes developed the mapping population. Kwabena Darkwa, Paterne Agre and Kolade Olufisayo executed the field phenotyping with contribution from Asrat Asfaw, Lava Kumar and Bunmi Olasanmi; Kwabena Darkwa performed some of the data analysis; Pierre Mournet

managed the genotyping for the sequencing; Paterne Agre managed the genotypic data and performed the QTL analysis; Paterne Agre and Kwabena Darkwa wrote the first draft with contributions from Asrat Asfaw and Bunmi Olasanmi. All authors reviewed and contributed to the final version.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Asiedu R, Sartie A. Crops that feed the world 1. Yams. *Food Secur.* 2010; 2(4):305-15.
2. Scarcelli N, Cubry P, Akakpo R, Thuillet AC, Obidiegwu J, Baco MN, Otoo E, Sonké B, Dansi A, Djedatin G, Mariac C. Yam genomics supports West Africa as a major cradle of crop domestication. *Science advances.* 2019; 5(5):1-7.
3. Darkwa K, Olasanmi B, Asiedu R, Asfaw A. Review of empirical and emerging breeding methods and tools for yam (*Dioscorea* spp.) improvement: Status and prospects. *Plant Breed.* 2019; 1-24.
4. Obidiegwu JE, Akpabio EM. The geography of yam cultivation in southern Nigeria: Exploring its social meanings and cultural functions. *J. Ethn. Foods.* 2017; 4(1):28-35.
5. Mignouna H, Mank R, Ellis T, Van Den Bosch N, Asiedu R, Abang M, Peleman J. A genetic linkage map of water yam (*Dioscorea alata* L.) based on AFLP markers and QTL analysis for anthracnose resistance. *Theor Appl Genet.* 2002; 105(5):726-35.
6. Baah FD, Maziya-Dixon B, Asiedu R, Oduro I, Ellis WO. Nutritional and biochemical composition of *D. alata* (*Dioscorea* spp.) tubers. *J Food Agric Environ.* 2009; 7(2): 373 – 378.
7. Abang MM, Fagbola O, Smalla K, Winter S. Two genetically distinct populations of *Colletotrichum gloeosporioides* Penz. causing anthracnose disease of yam (*Dioscorea* spp.). *J Phytopathol.* 2005; 153(3):137-42.
8. Abang MM, Green KR, Wanyera NW, Iloba C. Characterization of *Colletotrichum gloeosporioides* Penz. from yam (*Dioscorea* spp.) in Nigeria. In *Root crops in the 21st century. Proceedings of the 7th Triennial Symposium of the International Society for Tropical Root Crops–Africa Branch, Cotonou, Bénin 2001* (pp. 613-615).
9. Akem CN. Yam die-back and its principal cause in the yam belt of Nigeria. *Pak J. Biol. Sci.* 1999; 2:1106-9.
10. Selection of *alata* cultivars of low susceptibility to anthracnose (*Colletotrichum gloeosporioides*)
11. Winch JE, Newhook FJ, Jackson GV, Cole JS. Studies of *Colletotrichum gloeosporioides* disease on yam, *Dioscorea alata*, in Solomon Islands. *Plant pathol.* 1984; 33(4):467-77.
12. Mignucci J, Torres-López R, Hepperly P, Ramos-Busigo D, Campus M. Field diseases of tropical yams (*Dioscorea* spp.) and their control in Puerto Rico. In *VII th Symposium of the International Society for Tropical Root Crops.* 1988 (pp. 285-296).
13. Alleyne AT, O'Garro LW, Delauney AJ. Yam anthracnose in the English-speaking islands of the Eastern Caribbean–successes and research advances in disease management. *Trop Agric.* 2001; 75:53-7.

14. Frezal L, Jacqua G, Neema C. Study of the genetic diversity of *Colletotrichum gloeosporioides*, causative agent of anthracnose of white yam in Guadeloupe: In Proceedings of the 7th International Conference on Plant Diseases, 3–5 December 2003, Tours, France. 2003.
15. Abang MM, Hoffmann P, Winter S, Green KR, Wolf GA. Vegetative compatibility among isolates of *Colletotrichum gloeosporioides* from yam (*Dioscorea* spp.) in Nigeria. *J Phytopathol.* 2004; 152(1):21-7.
16. McDonald BA, Linde C. The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica.* 2002; 124(2):163-80.
17. Simons SA, Green KR. Epidemiology of yam anthracnose: Sources of inoculum. In: ProceedingS of 4th International Confeence of Plant Protection, Trop. Kuala Lumpur, Malaysia. 1994; 67-69.
18. Reis EM, Casa RT, Blum MM, dos Santos HP, Medeiros CA. Effects of cultural practices on the severity of leaf blotches of wheat and their relationship to the incidence of pathogenic fungi in the harvested seed. *Fitopatol Bras.* 1997; 22(3):407-12.
19. Mignucci JS, Hepperly PR, Green J, Torres-López R, Figueroa LA. Yam protection II. Anthracnose, yield, and profit of monocultures and interplantings. *Agric. Univ. P. R.* 1988; 72(2):179-89.
20. Palaniyandi SA, Yang SH, Cheng JH, Meng L, Suh JW. Biological control of anthracnose (*Colletotrichum gloeosporioides*) in yam by *Streptomyces* sp. MJM5763. *Appl. Microbiol.* 2011 Aug;111(2):443-55.
21. Hepperly P, Vazquez F. Resistance and scouting in the control of yam anthracnose of the winged yam (*Dioscorea alata*), In: Proceedings of 25th Annual Meeting of the Caribbean Food Crops Society Gosier, Guadeloupe. 1991; 587-596.
22. Bayart JD, Pallas B. Tolerance of yam anthracnose of a benzimidazole. Results of the first study conducted in Guadeloupe. *Phytoma.* 1994(461):37-40.
23. Asiedu R, Ng SY, Bai KV, Ekanayake IJ, Wanyera NM. Genetic improvement, In *Food yams: Advances in research*, IITA and NCRI, Nigeria. 1998:63-104.
24. Mignouna HD, Abang MM, Green KR, Asiedu R. Inheritance of resistance in water yam (*Dioscorea alata*) to anthracnose (*Colletotrichum gloeosporioides*). *Theor Appl Genet.* 2001; 103(1):52-5.
25. Petro D, Onyeka TJ, Etienne S, Rubens S. An intraspecific genetic map of water yam (*Dioscorea alata* L.) based on AFLP markers and QTL analysis for anthracnose resistance. *Euphytica.* 2011; 179(3):405-16.
26. Bhattacharjee R, Nwadili CO, Saski CA, Paterne A, Scheffler BE, Augusto J, Lopez-Montes A, Onyeka JT, Kumar PL, Bandyopadhyay R. An EST-SSR based genetic linkage map and identification of QTLs for anthracnose disease resistance in water yam (*Dioscorea alata* L.). *PloS one.* 2018; 13(10).
27. Cormier F, Lawac F, Maledon E, Gravillon MC, Nudol E, Mournet P, Vignes H, Arnau G. A reference high-density genetic map of greater yam (*Dioscorea alata* L.). *Theor and Appl Genet.* 2019; 132(6):1733-44.
28. Geffroy V, Sévignac M, De Oliveira JC, Fouilloux G, Skroch P, Thoquet P, Gepts P, Langin T, Dron M. Inheritance of partial resistance against *Colletotrichum lindemuthianum* in *Phaseolus vulgaris* and

- co-localization of quantitative trait loci with genes involved in specific resistance. *Mol Plant Microbe Interact.* 2000; 13(3):287-96.
29. Ling H, Zhao J, Zuo K, Qiu C, Yao H, Qin J, Sun X, Tang K. Isolation and expression analysis of a GDSL-like lipase gene from *Brassica napus* L. *J Biochem Mol Biol.* 2006; 39(3):297.
 30. Kwon SJ, Jin HC, Lee S, Nam MH, Chung JH, Kwon SI, Ryu CM, Park OK. GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in *Arabidopsis*. *Plant J.* 2009; 58(2):235-45.
 31. Hong JK, Choi HW, Hwang IS, Kim DS, Kim NH, Choi DS, Kim YJ, Hwang BK. Function of a novel GDSL-type pepper lipase gene, CaGLIP1, in disease susceptibility and abiotic stress tolerance. *Planta.* 2008; 227(3):539-58.
 32. Sun X, Cao Y, Yang Z, Xu C, Li X, Wang S, Zhang Q. Xa26, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. *Plant J.* 2004; 37(4):517-27.
 33. Llorente F, Alonso-Blanco C, Sánchez-Rodríguez C, Jorda L, Molina A. ERECTA receptor-like kinase and heterotrimeric G protein from *Arabidopsis* are required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. *Plant J.* 2005; 43(2):165-80.
 34. Mbengue M, Bourdais G, Gervasi F, Beck M, Zhou J, Spallek T, Bartels S, Boller T, Ueda T, Kuhn H, Robatzek S. Clathrin-dependent endocytosis is required for immunity mediated by pattern recognition receptor kinases. *Natl. Acad. Sci.* 2016; 113(39):11034-9.
 35. Ibraheem F, Gaffoor I, Chopra S. Flavonoid phytoalexin-dependent resistance to anthracnose leaf blight requires a functional yellow seed1 in *Sorghum bicolor*. *Genetics.* 2010; 184(4):915-26.
 36. Ibraheem F, Gaffoor I, Tan Q, Shyu CR, Chopra S. A sorghum MYB transcription factor induces 3-deoxyanthocyanidins and enhances resistance against leaf blights in maize. *Molecules.* 2015; 20(2):2388-404.
 37. van den Burg HA, Tsitsigiannis DI, Rowland O, Lo J, Rallapalli G, MacLean D, Takken FL, Jones JD. The F-box protein ACRE189/ACIF1 regulates cell death and defense responses activated during pathogen recognition in tobacco and tomato. *The Plant Cell.* 2008; 20(3):697-719.
 38. Takemoto D, Rafiqi M, Hurley U, Lawrence GJ, Bernoux M, Hardham AR, Ellis JG, Dodds PN, Jones DA. N-terminal motifs in some plant disease resistance proteins function in membrane attachment and contribute to disease resistance. *Mol Plant Microbe Interact.* 2012; 25(3):379-92.
 39. Asfaw A, Standard Operating Protocol for Yam Variety Performance Evaluation Trial, IITA, Ibadan, Nigeria. 2016; 1 - 27.
 40. Campbell CL, Madden LV. *Introduction to plant disease epidemiology*, John Wiley and Sons, New York. 1990.
 41. Elshire RJ, 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).
 42. R Core Team. *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. 2017. Available at <https://www.R-project.org/>

43. Scarcelli N, Mariac C, Couvreur TL, Faye A, Richard D, Sabot F, Berthouly-Salazar C, Vigouroux Y. Intra-individual polymorphism in chloroplasts from NGS data: Where does it come from and how to handle it?. *Mol Ecol Resour.* 2016; 16(2):434-45.
44. Martin M, Cutadapt removes adapter sequences from high-throughput sequencing reads *EMBnet J.* 2011; (17):10 - 12.
45. Sarah G, Homa F, Pointet S, Contreras S, Sabot F, Nabholz B, Santoni S, Sauné L, Ardisson M, Chantret N, Sauvage C. A large set of 26 new reference transcriptomes dedicated to comparative population genomics in crops and wild relatives. *Mol Ecol Resour.* 2017; 17(3):565-80.
46. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G. The variant call format and VCFtools. *Bioinformatics.* 2011; 27(15):2156-8.
47. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81(3):559-75.
48. Mollinari M, Olukolu BA, Pereira GD, Khan A, Gemenet D, Yenchu GC, Zeng ZB. Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. *G3-GENES GENOM GENET.* 2020; 10(1):281-92.
49. Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen S, Yandell BS, Churchill GA. R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics.* 2019; 211(2):495-502.
50. Ouellette LA, Reid RW, Blanchard SG, Brouwer CR. LinkageMapView—rendering high-resolution linkage and QTL maps. *Bioinformatics.* 2018; 34(2):306-7.
51. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics.* 1994; 138(3):963-71.

Supplementary Figure Legends

Supplementary Figure 1: Recombination fraction of genotyped GBS markers. A pairwise comparison of recombination fractions (upper right triangle) and LOD score (lower right triangle) across the 20 chromosomes.

Supplementary Figure 2: Allocation of the 3257 SNP markers into 20 different linkage genetic groups

Figures

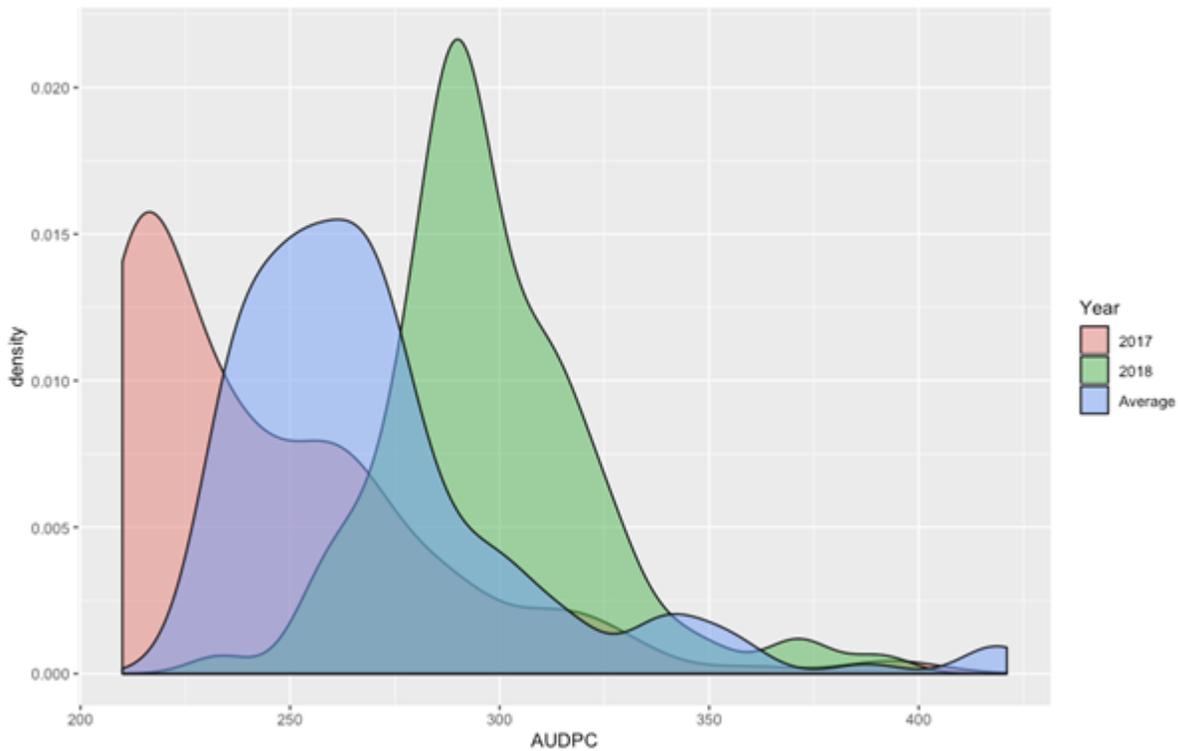


Figure 1

Frequency distribution showing the reaction of the mapping population to yam anthracnose disease for individual years (2017 and 2018) and average across two years.

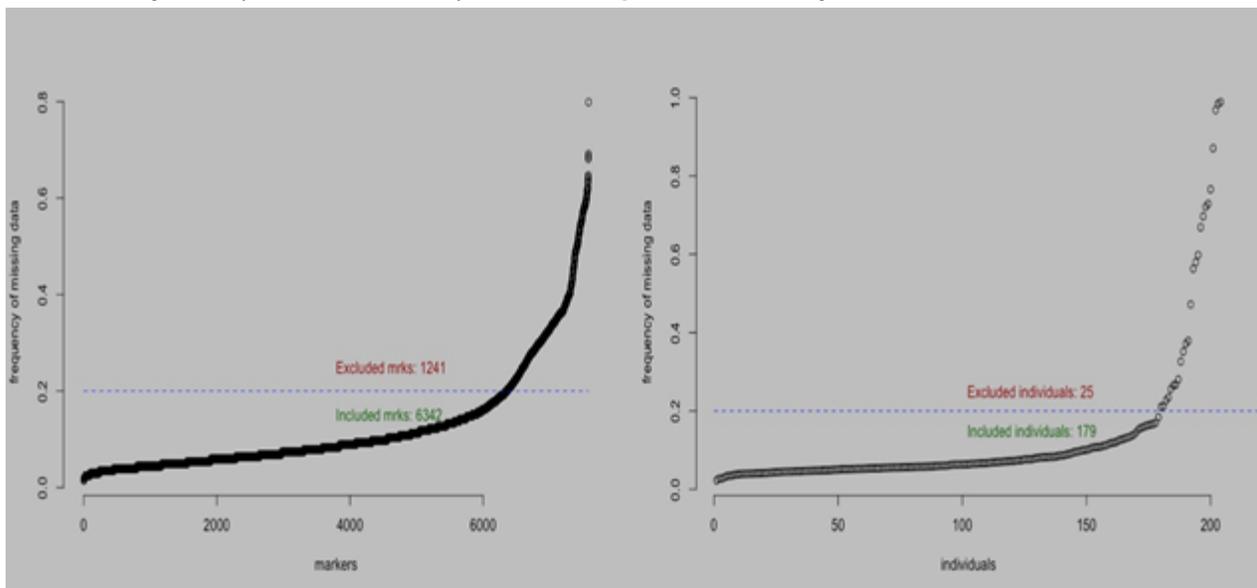


Figure 2

SNP quality assessment, the left Figure shows the informative markers after Chi-square test (Missing threshold = 20%) while the Figure on the right indicates the accessions with good coverage. The blue horizontal dotted lines represent the threshold.

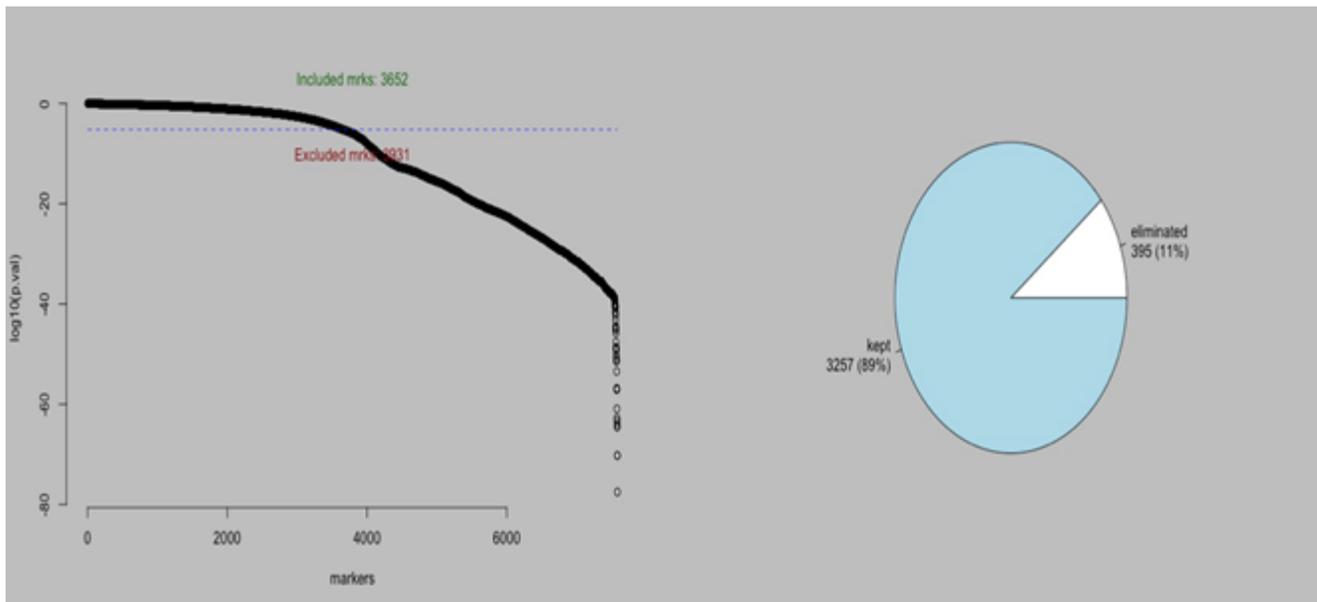


Figure 3

Assessment of SNP quality through conservative Bonferroni method

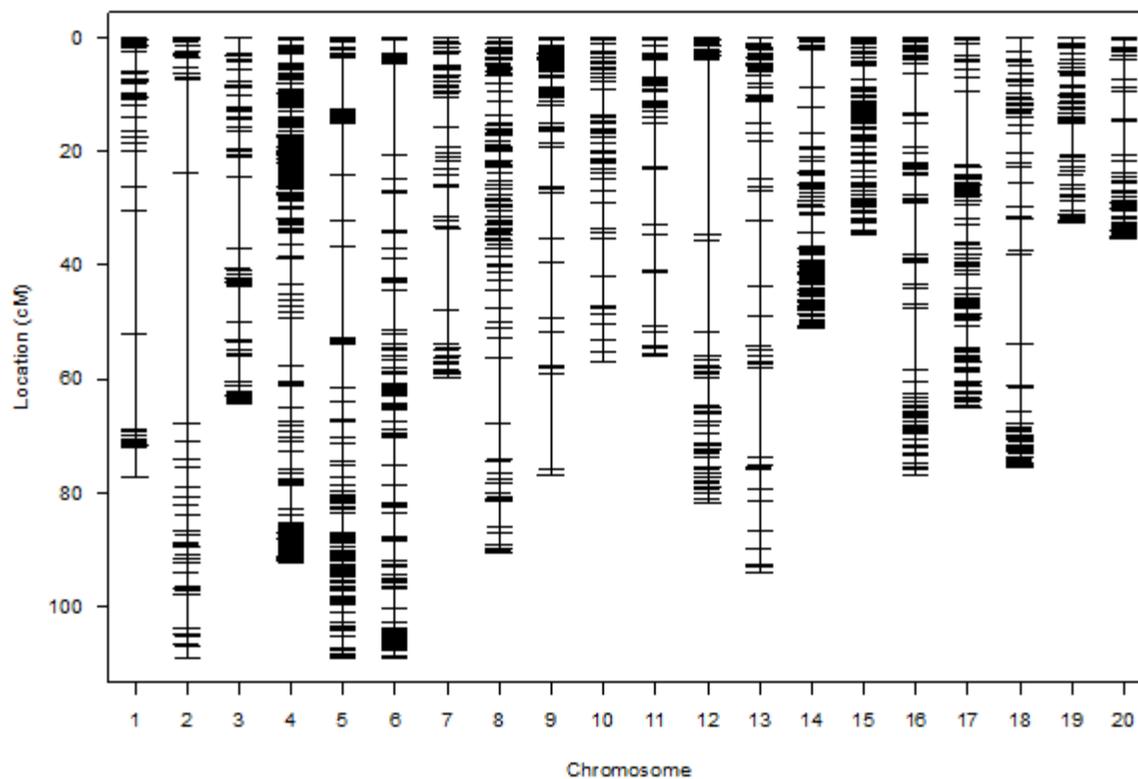


Figure 4

Genetic linkage map showing the SNP distribution across the 20 chromosomes. Each vertical line represents a yam chromosome with genetic distance in Kosambi units (cM).

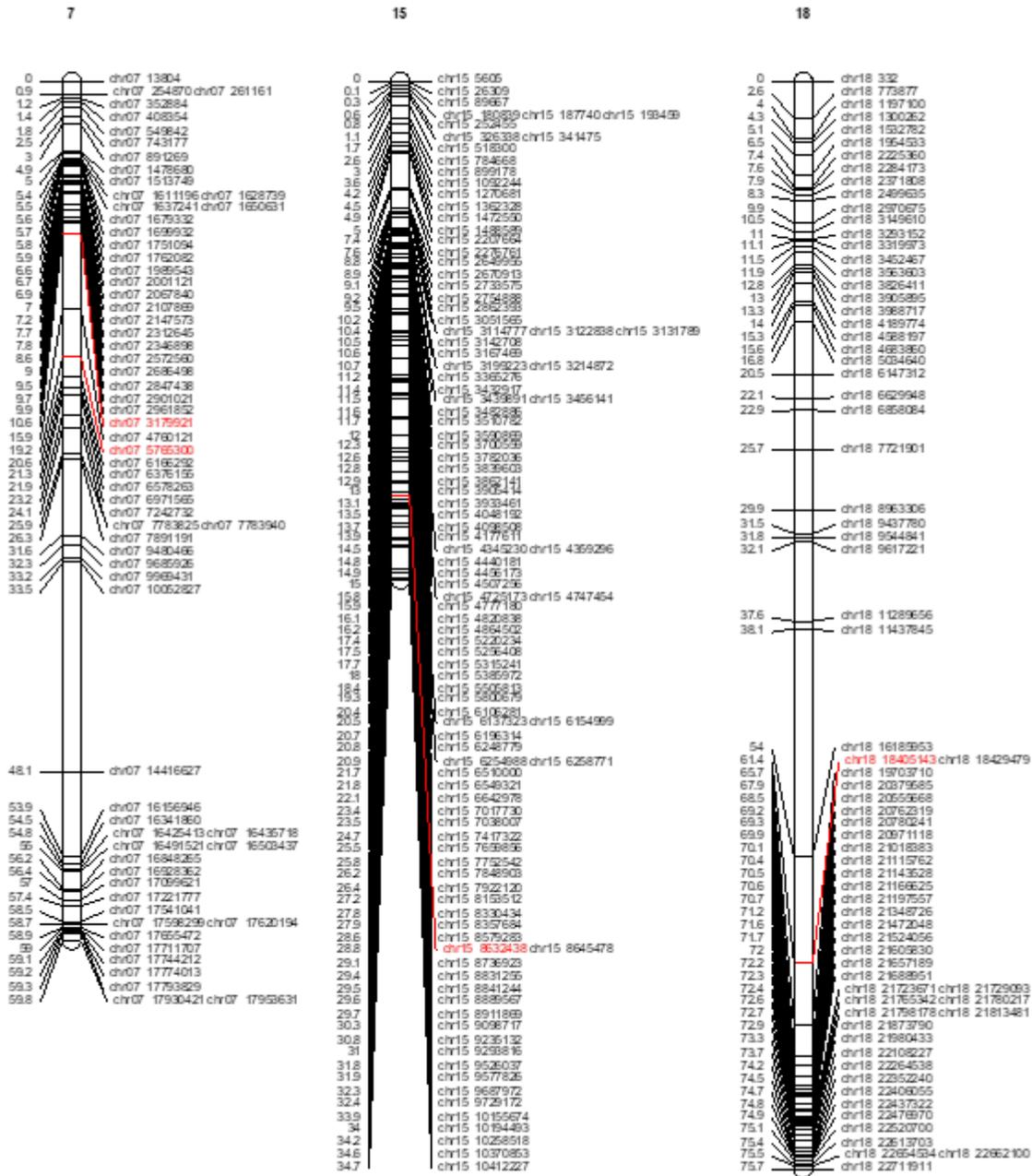


Figure 5

Genetic map of water yam showing significant QTLs associated with yam anthracnose disease resistance. Only those chromosomes where significant QTL were located are shown. The identified QTLs are highlighted in red on each chromosome

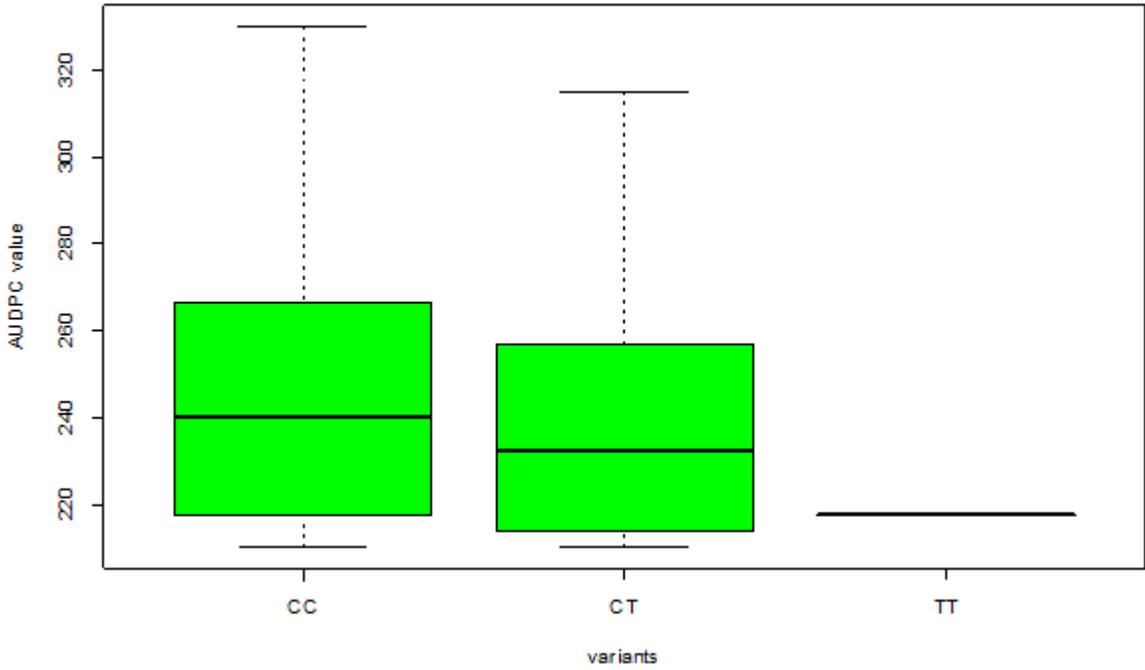


Figure 6

The boxplot shows the effect of the different alleles (variants) of chr7_3179921 on the AUDPC estimates. The letters on the X axis represent alleles (CC, CT/TC and TT).

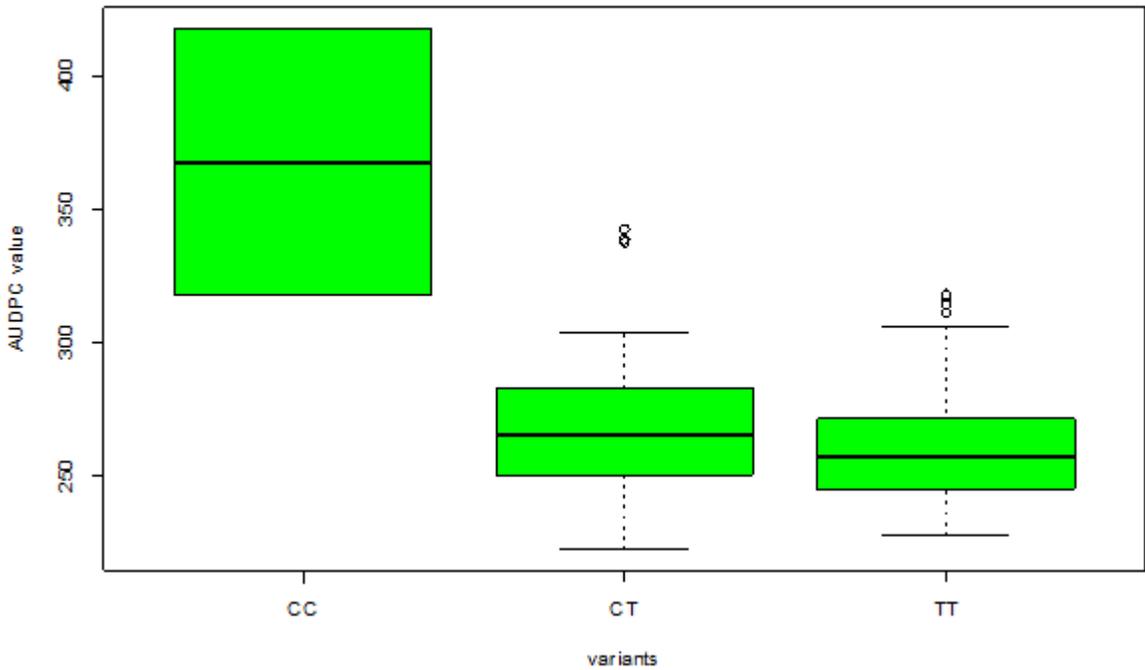


Figure 7

The boxplot shows the effect of the different alleles (variants) of chr07_5765300 on the AUDPC estimates. The letters on the X axis represent alleles (CC, CT/TC and TT).

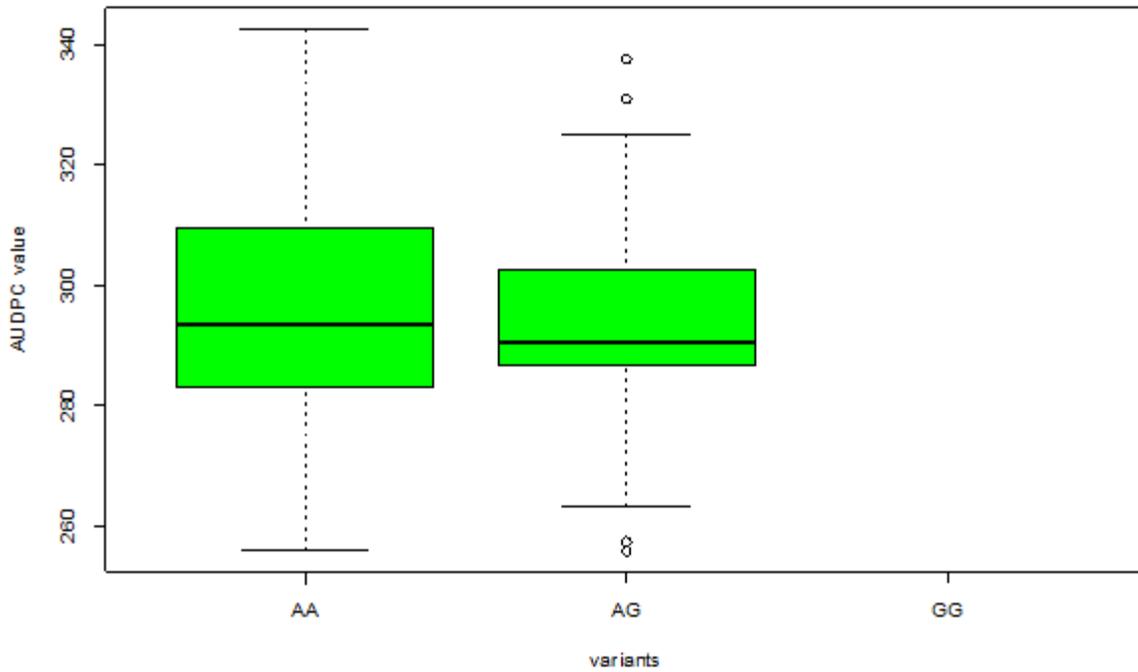


Figure 8

The boxplot shows the effect of the different alleles (variants) of chr15_8632438 on the AUDPC estimates. The letters on the X axis represent alleles (AA, AG/GA and GG).

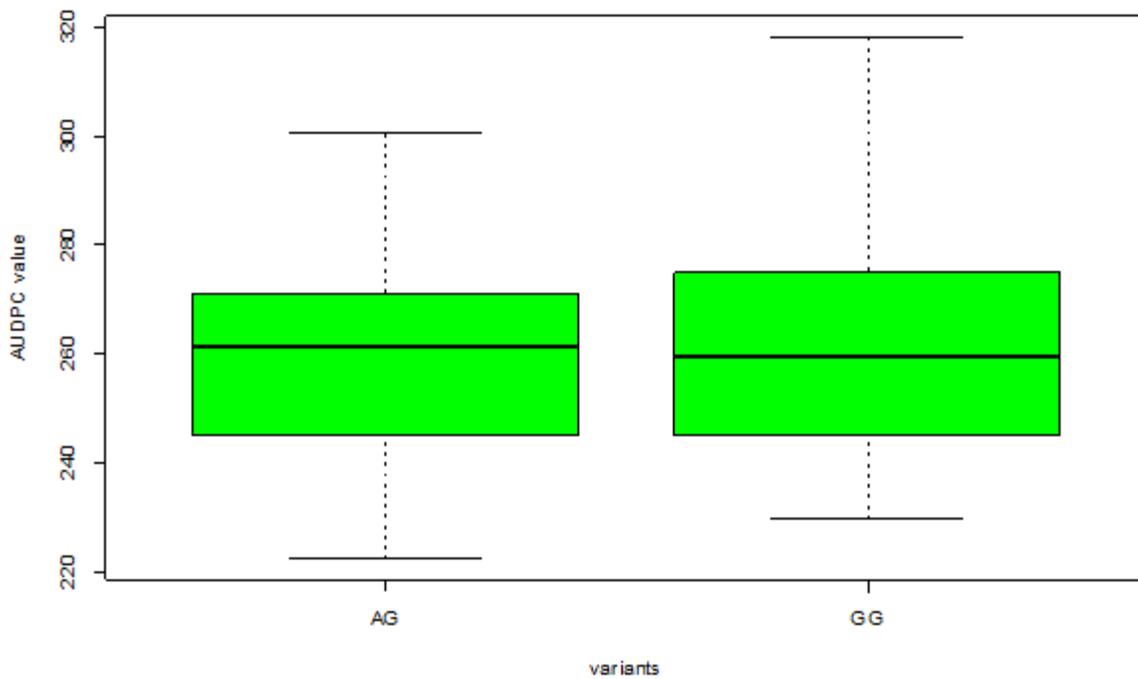


Figure 9

The boxplot shows the effect of the different alleles (variants) of chr18_18405143 on the AUDPC estimates. The letters on the X axis represent alleles (AA, AG/GA and GG).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigures.pdf](#)
- [Equations.pdf](#)