ORIGINAL ARTICLE

SNP markers associated with resistance to frosty pod and black pod rot diseases in an F$_1$ population of *Theobroma cacao* L.

Osman A. Gutiérrez[1] · Alina S. Puig[1] · Wilbert Phillips-Mora[2] · Bryan A. Bailey[3] · Shahin S. Ali[4] ·
Keithanne Mockaitis[5] · Raymond J. Schnell[6] · Donald Livingstone[6] · Guiliana Mustiga[6] ·   Stefan
 Royaert[6] · Juan Carlos Motamayor[6]

Abstract

Economically, cacao (*Theobroma cacao* L.) is a major tropical commodity for the Americas; however, severe losses due to *Moniliophthora roreri* (Cif. and Par.), which causes frosty pod rot (FPR), and *Phytophthora* spp., which causes black pod rot (BPR), have reduced cacao production in the Americas. The objectives of this study are to (i) re-confirm the QTL using different marker set; (ii) discover new QTL associated with FPR and BPR resistance using SNP markers; and (iii) find genes in the candidate QTL regions. At CATIE in Turrialba, Costa Rica, an F$_1$ mapping population of cacao was obtained by crossing "POUND 7," a clone moderately susceptible to FPR and resistant to BPR, with "UF 273," resistant to FPR and highly susceptible to BPR. A total of 179 F$_1$ progeny were fingerprinted with 5149 SNP markers and a dense linkage map composed of 10 linkage groups was developed using 2910 polymorphic SNP markers. Also segregating F$_1$ trees were screened for resistance to FPR and BPR diseases. Seven QTL previously reported on chromosomes 2, 7, and 8 for FPR resistance and on chromosomes 4, 8, and 10 for BPR resistance were confirmed. Additionally, eight QTL were identified for FPR resistance (chromosomes 4, 9, and 10) and BPR resistance (chromosome 2). The expression of genes commonly associated with plant defense and disease resistance that are located within the identified QTL was confirmed.

Keywords Linkage map · Matina 1-6 v1.1 · Composite interval mapping · *Moniliophthora roreri* · *Phytophthora palmivora* · Single nucleotide polymorphisms · Quantitative trait loci · Disease resistance

[1] USDA-ARS Subtropical Horticulture Research Station, 13601 Old Cutler Road, Miami, FL 33158, USA

[2] CATIE, Tropical Agricultural Research and Higher Education Center, Turrialba 7170, Costa Rica

[3] USDA-ARS Sustainable Perennial Crops Laboratory, 10300 Baltimore Avenue, Bldg. 001, BARC-West, Beltsville, MD 20705, USA

[4] Department of Viticulture and Enology, University of California, Davis, CA 95616, USA

[5] Pervasive Technology Institute, Indiana University, 2709 E. 10th St, Bloomington, IN 47408, USA

[6] MARS, Inc., 13601 Old Cutler Road, Miami 33158, FL, USA

Introduction

Cacao (*Theobroma cacao* L.) is an evergreen tree indigenous to the Amazon basin. It can be found growing in many countries between latitudes 20° S and 20° N (Cope 1984). Fermented and dried cacao seeds (beans) are used by the confectionary industry, as the main ingredient in chocolate, as well as in the food and beverage, cosmetic, and pharmaceutical industries.

It was thought that cacao was domesticated in Mesoamerica by the Olmec, who transferred their knowledge to the Toltec, Mayan, and Aztec civilizations (Bartley 2005; Coe and Coe 2013). However, recent findings by Zarrillo et al. (2018) indicated that the cacao center of diversity was located on the upper Amazon region of northwest South America, which was also the center of domestication. *Theobroma cacao* is a diploid species (2n = 2x = 20) within the family Malvaceae (Whitlock and Baum 1999), initially divided into three major morphological/geographic groups: Criollo, Forastero, and

Trinitario (Cheesman 1944). After analyzing 1241 diverse accessions with 106 simple sequence repeat (SSR) markers, Motamayor et al. (2008) later proposed classifying cacao germplasm into 10 major groups: Amelonado, Contamana, Criollo, Curaray, Guiana, Iquitos, Marañon, Nacional, Nanay, and Purús.

Currently, cacao production is concentrated in West Africa (72.3%); however, the Americas (18.3%) and Asia (9.4%) also contribute to the world supply of cocoa (ICCO 2017). Disease and insect pests contribute to losses of about 30% worldwide (Ploetz 2016). Two of the diseases responsible for the greatest losses in cacao are frosty pod rot (FPR), caused by *Moniliophthora roreri*, and black pod rot (BPR), caused by several different *Phytophthora* species (*P. capsici*, *P. tropicalis*, *P. citrophthora*, *P. megakarya*, and *P. palmivora*). BPR pathogens are present in all cacao producing countries and cause annual yield losses of 20 to 30% (Surujdeo-Maharaj et al. 2016). In contrast, FPR causes less overall reduction in world cacao production due to its limited geographic distribution. However, in some regions, average pod losses from 30% to above 90% have been reported (Phillips-Mora and Wilkinson 2007; Bailey et al. 2018). FPR is present in most cacao growing areas of the Americas, excluding Brazil and all Caribbean countries except Jamaica where it was reported in 2016 (Johnson et al. 2017). In places where both BPR and FPR diseases co-occur, FPR causes the highest yield losses (Evans 2007; Phillips-Mora and Wilkinson 2007). It is expected that FPR will continue to disperse across the cacao producing regions of the Western Hemisphere (Bailey et al. 2018).

BPR resistance is estimated by measuring the diameter of lesions developing on pods (Phillips-Mora and Castillo 1999), by estimating lesion area (Campbell et al. 2015), or by determining the area under the disease-progress curve (Ling et al. 2017). Importantly, screening methodology impacts the outcome of disease resistance studies. Greater susceptibility to BPR was found, for example, when inoculations were applied to detached pods versus attached pods (Iwaro et al. 1998). The inoculation of wounded pods vs. non-wounded pods has yielded different results, as these methods screen for pre- and post-penetration resistance, respectively (Surujdeo-Maharaj et al. 2016).

In addition to pods, since the early 1950s, leaf inoculations have been used to screen for BPR resistance, and leaf disk inoculation methodology has been standardized (Holliday 1954; Nyassé et al. 1995; Akaza et al. 2009).

QTL analyses have been conducted to determine which regions of the genome can be linked to resistance to *Phytophthora* spp. (Crouzillat et al. 2000; Brown et al. 2007; Motilal et al. 2016). Since different screening methodologies were employed in these discoveries, it is possible that the QTL identified are associated to distinct mechanisms of resistance.

For example, QTL detected using leaf disk inoculation may not correspond to QTL identified using artificial pod inoculation methodology or field resistance (Lanaud et al. 1999; Flament et al. 2001). In addition, mapped QTL for BPR resistance based on field incidence and leaf disk inoculation experiments has also provided different outcomes. Akaza et al. (2016) using the following crosses, ("SCA 6" × "H") x "C1," ("POUND 7" × "ICS 100") × "C1," and ("POUND 7" × "ICS 95") × "C1," reported that the leaf disk inoculation methodology led to the identification of QTL on linkage groups (LGs) 1, 3, and 10 from the parent "POUND 7" × "ICS 100," while incidence of BPR in the field led to the identification of disease-resistance associated QTL in all three of the parents ("POUND 7" × "ICS 95," LGs 2 and 4; "SCA 6" × "H," LGs 1, 6, and 8; "POUND 7" × "ICS 100," LGs 4 and 6). Only chromosome 1 carried QTL for resistance to *P. palmivora* based on both screening techniques (Akaza et al. 2016). Contradictory results may also arise due to the variation in the virulence among isolates used in each region or due to parental off-types (Cervantes-Martinez et al. 2006; Olasupo et al. 2018).

While resistance to BPR has been identified in several germplasm accessions, resistance to FPR is relatively uncommon. Using seventy new cacao clones, Phillips-Mora and Castillo (1999) conducted a screening experiment at the Centro Agronómico Tropical de Investigación y Enseñanza (CATIE) in Turrialba, Costa Rica. They found that only 2 clones (3%) were characterized as moderately resistant (MR). Furthermore, they stated that based on their screening results, only 10 clones (2.3%) out of 441 clones that represent 56% of the CATIE collection were identified as resistant or moderately resistant to FPR. In the latest report, Phillips-Mora et al. (2017) stated that out of 1260 clones from the CATIE collection, 76 (6%) showed tolerance to FPR. This recent result confirms the low number of accessions with FPR tolerance/resistance. The levels of internal severity (IS), external severity (ES), and disease incidence were assessed following artificial inoculation (Phillips-Mora 1996; Suárez-Capello 1999), as well as measuring disease incidence under natural infection conditions (Galindo and Enriquez 1984). Bejarano (1961) successfully obtained *M. roreri* infection by spraying pods with a spore suspension and containing them within a polyethylene bag. He demonstrated that wounds were not necessary for infection to develop. Since then, this screening methodology has been modified numerous times and optimized to differentiate levels of resistance among genotypes (Sotomayor 1965; Sánchez 1982; Phillips-Mora and Galindo 1989; Sánchez and Gonzalez 1989).

An important goal in cacao improvement programs around the world is breeding for disease resistance; however, only 30% of material currently being cultivated is made up of improved varieties (Gutiérrez et al. 2016). Breeding disease-resistant plants involves crossing and selecting genotypes in

order to accumulate desirable genes for resistance and agronomic traits (Keane 2012). Since 1999, the USDA-ARS has been collaborating with CATIE and MARS, Inc., on a genomic-assisted breeding program with an emphasis on disease resistance (Schnell et al. 2005; Schnell et al. 2007). This program uses genomic-assisted breeding methodologies to accelerate the breeding process. Developing disease-resistant germplasm with high yields would provide a stable supply of raw material to the worldwide chocolate confectionary industry.

Brown et al. (2007) in an F_1 population of the cross "POUND 7" × "UF 273" determined that there were five QTL for FPR resistance. They used 180 markers (single sequence repeat, resistance gene homolog sequences, and one "WRKY" stress-related marker) to build a linkage map containing 169 loci and identified putative QTL for resistance to both FPR and BPR. The five QTL associated with resistance to FPR were located on linkage groups 2, 7, and 8, and three QTL associated with BPR resistance on chromosomes 4, 8, and 10.

The current study examines the same mapping population used by Brown et al. (2007), genotyped with far more markers (5149 SNPs versus the original 180 SSRs) using an 6K Illumina SNP chip developed by Livingstone et al. (2015), resulting in a more dense genetic linkage map as well as more precise QTL identification. The purpose of this study is to (i) re-confirm the QTL using a different marker set; (ii) discover new QTL associated with FPR and BPR resistance using SNP markers; and (iii) find genes in the candidate QTL regions.

Materials and methods

Plant materials

A heterozygous F_1 mapping population was developed at CATIE by crossing "POUND 7," a clone moderately susceptible to FPR and resistant to BPR, with "UF 273," which is resistant to FPR and highly susceptible to BPR (Phillips-Mora et al. 2013). This population segregates for disease resistance traits, making it ideal for use in mapping and QTL identification. One of the 5 "UF 273" accessions used as pollen donors was found to be an off-type and is referred to as "UF273" Type II. It differed at 22 of the 180 SSR alleles tested by Brown et al. (2007), and individuals of the F_1 population were designated Type I ($n = 185$) or Type II ($n = 71$) to indicate from which father they arose (Cervantes-Martinez et al. 2006). In addition, after screening Type I and Type II progenies with 5149 SNP markers, Livingstone et al. (2015) obtained 0.162 as the proportion of SNP loci that were different between the two parental accessions, "UF 273" Type I and "UF 273" Type II. Further disease resistance screening of these trees concluded that the "UF 273" Type I tree was resistant to FPR while the

"UF 273" Type II was classified as moderately resistant (Phillips-Mora et al. 2013). The original F_1 population was comprised of 256 individuals; however, only the subset of 179 Type I F_1 individuals was included in this study because they were descendants of the "UF 273" Type I parent.

Trees were planted in 1998 at the CATIE farm, La Montaña. Planting sites and plot arrangements are described in detail in Brown et al. (2007). Artificial inoculations of pods were carried out on the complete population [Type I ($n = 185$) and Type II ($n = 71$)] from 2000 to 2004 to assess resistance to FPR and BPR by Brown et al. (2007). However, only data from Type I individuals were considered in this research.

Genomic DNA isolation and SNP genotyping

DNA extraction was conducted using the methodology employed by Livingstone et al. (2015).

For each sample, 300 mg of leaf material was homogenized in liquid nitrogen, suspended in 1 mL wash buffer (100 mM Hepes, 0.1% PVP 40 (w/v), and 4% 2-mercapto-ethanol), vortexed briefly, and centrifuged (8400 ×g for 5 min). Wash buffer was removed from the pelleted material, and 1 mL of fresh wash buffer applied followed by vortexing and centrifugation (8400 ×g for 5 min); this process was repeated 4–6 times until the supernatant was no longer viscous. Nuclei extraction buffer (15% sucrose, 50 mM Tris–HCl, 50 mM EDTA pH 8.0, and 500 mM NaCl) was added to the pellet and it was vortexed until the pellet had been resuspended back into solution. Once resuspended, the sample was incubated (50 °C) for 15 min with gentle vortexing every 2 min. After incubation, the sample was centrifuged (8400 ×g for 5 min.), the supernatant removed, and the pellet resuspended in 450 μL of buffer (20 mM Tris–HCl with 10 mM EDTA) and 80 μL of 10% SDS. The sample was subjected to another 15-min incubation period, but this time at 70 °C. After incubation, the sample was cooled to room temperature and subsequently, 300 μL of ice-cold, −20 °C, 7.5 M NH4OAc was added. The sample was placed on ice for 30 min and then the tube was centrifuged at top speed for 15 min at 4 °C. The aqueous top layer was reserved in a new tube, and an equal part of isopropanol was added, followed by a centrifugation step (max speed, 15 min) at 4°C. Ice-cold 70% ETOH was used to wash the DNA pellet twice after which it was resuspended in 100 μL Tris–HCl (10 mM). Leaf material was obtained from the parents "POUND 7" and "UF 273" Type I and "UF 273" Type II, as well as from the F_1 progenies and the extracted DNA was submitted to Illumina for genotyping using the 6K Illumina SNP chip developed by Livingstone et al. (2015). SNP ID, mapping population, position, linkage group location, gene model, and sequences are provided in supplemental file 3 of Livingstone et al. (2015).

Disease resistance screening

Frosty pod rot

In order to measure resistance to FPR, fruits from 25–35 selected trees were inoculated with *M. roreri* during an inoculation event, which refers to inoculations performed within the same day, using the method designed by Phillips-Mora (1996). Young attached pods, 2–3-month-old, were artificially inoculated by spraying a spore suspension of *M. roreri*, covering them with a bag for 2 days to maintain humidity, then evaluating disease severity 9 weeks after inoculation. For each inoculation date, during the 4-year period, between 1 and 10 available pods were sampled per selected tree. Each genotype was inoculated between 3 and 17 times, depending on fruit availability. Nearly all available pods were inoculated in each event.

FPR disease ratings are based on ES and IS which are measured 9 weeks after inoculation. ES quantifies symptom development on the pod surface using a five-point scale (Sánchez et al. 1987; Sánchez and Gonzalez 1989), where: 0 = healthy pod; 1 = water soaking; 2 = swellings and/or premature ripening; 3 = necrosis; 4 = mycelium covering less than 25% of the necrotic area; and 5 = mycelium covering more than 25% of the necrotic area. IS, based on the percentage of internal necrosis, was rated on a scale of 0 to 5, where: 0 = 0%; 1 = 1–20%; 2 = 21–40%; 3 = 41–60%; 4 = 61–80%; and 5 = >80% (Sánchez et al. 1987). For each genotype, an average rating was calculated for all pods within a single inoculation event (Gutiérrez et al. 2016). Because the production of pods was not uniform among all Type I F_1 individuals during the different inoculation events, phenotypic data was retained only for 169 Type I F_1 progenies. A minimum of three separate inoculation events, including at least five pods per tree, was required for progenies to be included in the analysis.

Black pod rot

Phytophthora palmivora inoculations were done by spraying pods with zoospore suspensions and covering them with a plastic bag to maintain the high humidity conditions required for infection. Between 1 and 13 pods from each tree were sampled at each inoculation date, depending on availability. Genotypes with fewer than three inoculations were excluded. The number of inoculations per genotype ranged from 4 to 19.

Disease severity was quantified as average diameter (cm) of the largest lesion on each pod 10 days after inoculation (DL10) (Crouzillat et al. 2000). All pods from a single inoculation event were averaged to give a mean rating per tree. Only 175 Type I F_1 trees were retained after using the same criteria previously described for the FPR resistance evaluation

analysis regarding a minimum number of pods as well as inoculation events.

Data analysis

Analyses for resistance to FPR included genotypes with at least five total inoculated fruits and three inoculation events. Distribution of ES and IS for each genotype was assessed using Proc Univariate in SAS (SAS Institute 2016) and found to be normal. However, variances were found to be unequal based on Levene's test for homogeneity. Internal and external disease severity values were analyzed using a Welch ANOVA, which adjusts for unequal variance, with SAS V 9.4 (SAS Institute 2016), and least square means were calculated for each tree. Very small quantities (0.001–0.003) were added to observations within a given genotype when they shared identical values for the model to converge.

Analyses for resistance to BPR included genotypes with a minimum of four inoculation events. Three genotypes were excluded, as they had two or fewer inoculation events. Normal distribution of DL10 was confirmed as described above; however, variances were again found to be unequal based on Levene's test for homogeneity. A generalized mixed linear model was run on DL10 to calculate least square means and adjust for unequal variance.

JoinMap® 5 (Van Ooijen 2018) was utilized to assemble genetic linkage maps using SNP data from 185 F_1 progenies and cross-pollinating (CP) population data classification scheme. Linkage map computations were performed using SNP genotyping data that was converted to JoinMap format segregation patterns <*lm* × *ll*> (maternal parent), <*nn* × *np*> (paternal parent), and <*hk* × *hk*> (both parents) using an R script program. Distorted loci that did not follow the expected segregation ratios of 1:1 for the *lm* × *ll* and *nn* × *np* segregation patterns as well as 1:2:1 *hk* × *hk* segregation pattern were discarded based on a chi-square test (P < 0.01). Later, the segregation patterns <*lm* × *ll*> and <*hk* × *hk*> were used to construct the maternal parent ("POUND 7") linkage map. In contrast, <*nn* × *np*> and <*hk* × *hk*> were utilized for the paternal progenitor ("UF 273" Type I). Linkage grouping was performed using the independence LOD option and 4.0 was the minimum LOD value used to select the linkage groups. The multipoint frequency recombination estimation in the maximum likelihood mapping procedure was estimated by a quicker deterministic EM algorithm (Van Ooijen 2018) and was also utilized to construct the linkage maps. The Haldane mapping function (Haldane 1919) was used to calculate mapping distances. To select the best locus order in each individual linkage group, map calculations were performed ten times and the best maps were selected based on the log-likelihood criterion (LogE-likelihood) that consists of choosing the maps with the least negative log-likelihood value (Van Ooijen and Jansen 2013). In addition, nearest–neighbor

stress test (N.N. Stress) was used to observe the quality of the simulated annealing using a threshold value of 2. Identical loci were identified by JoinMap® 5 based on identical segregation patterns and linkage map computations were only performed using representative loci from each linkage bin having the minimum missing genotype rate. Linkage map and location of the QTL on the genetic map were created using LinkageMapView software version 2.1.2 (Ouellette et al. 2018).

QTL analyses were performed using a mixed model approach to the genome-wide scan with GenStat® 19.0 software (VSN International 2017) in the following order. First, a calculation of the genetic predictors was conducted followed by an initial genome-wide scan using marker regression and simple interval mapping (SIM) methodology (Lander and Botstein 1989) to find candidate QTL positions. Subsequently, the selected SNP markers associated with QTL for FPR and BPR resistance were used as co-factors in several rounds of composite interval mapping (CIM) (Jansen and Stam 1994; Zeng 1994). Lastly, a final multi-QTL model developed by Boer et al. (2017) and described below was fit using back-selection and the CIM's set of candidates QTL to obtain the final set of estimated QTL effects.

$$y_i = \mu + \alpha^{a1} x_i^{a1} + \alpha^{a2} x_i^{a2} + \alpha^d x_i^d + e_i$$

where:

y_i is the trait mean for genotype i
μ is the overall mean
α^{a1} is the QTL additive effect of the first parent (POUND 7) at the position being examined
x_i^{a1} is the additive genetic predictor for parent 1 allele inherited by progeny i at the position being examined
α^{a2} is the QTL additive effect of the second parent (UF 273) at the position being examined
x_i^{a2} is the additive genetic predictor for parent 2 allele inherited by progeny i at the position being examined
α^d is the dominance QTL effect at the position being examined
x_i^d is the dominance genetic predictor for genotype i at the position being examined
e_i is the residual effect

QTL were treated as fixed effects and their significance in the model was determined by the Wald test. Calculation of genetic predictors, SIM, CIM, and final QTL model were performed using the QIBDPROBABILITIES, QSQTLSCAN, QCANDIDATES, and QSBACKSEL routines and the CP population option available in GenStat®.

At each QTL, four genotypic classes are inferred based on inheritance of haplotype "a" or "b" from POUND 7 and haplotype "c" or "d" from UF 273. The haplotypes represent identity-by-descent at the locus inferred from the linkage map and surrounding markers, rather than identity-in-state at a single marker. The means for the *ac*, *ad*, *bc*, and *bd* genotypes were then used to estimate the additive and dominance effects at the QTL. The population structure permits comparison of the average difference between the two alleles inherited from Pound 7 (a_1 = gp_additive), the two alleles inherited from UF 273 (a_2 = gp_additive2), and a within-locus allelic interaction effect (d = gp_dominance). This interaction effect is referred to as a dominance effect, but it is not a classical dominance effect representing the difference between the heterozygote and mid-parent values at a biallelic locus, because none of the progenies are homozygous for a parental haplotype defined by identity-by-descent. Instead, this dominance effect represents the deviation of specific combinations of alleles from the two parents compared to their expected genotypic values based on the average allelic effects. Additive and dominance effects are related to the genotypic class means as follows:

$$ac : mu - a_1 - a_2 + d$$
$$ad : mu - a_1 + a_2 - d$$
$$bc : mu + a_1 - a_2 - d$$
$$bd : mu + a_1 + a_2 + d$$

Also the percentages of the variance explained by an SNP marker were calculated using the QIBDPROBABILITIES procedure's estimate effects option on the final QTL model and the mixed model REML procedure and VCOMPONENTS directive (Payne et al. 2015; Boer et al. 2017). The threshold value [−Log10(P) ($P < 0.05$), ($P < 0.01$), and ($P < 0.001$)], used to determine the significance of a QTL, was estimated by GenStat® using a modified Bonferroni correction base, a methodology developed by Li and Ji (2005). SNP markers associated with top LOD values were designated as QTL linked to SNPs. After the QTL analysis was completed, parental and progeny haplotypes of the SNP markers and flanking markers associated with QTL for each one of the evaluated traits were obtained and phased with iXora (Utro et al. 2013) and JoinMap® 5 (Van Ooijen 2018). Since the QTL model is based on inferring inheritance of parent 1 alleles (a or b) and parent 2 alleles (c or d) using linkage information and local multi-locus genotypes, optimal genotypic classes are identified in terms of these inferred identity-by-descent allelic representations. To identify specific combinations of SNP alleles very closely linked to the QTL peak positions that represent these genotypic classes, pairs of SNPs surrounding each QTL with alternate segregation patterns were used to define the genotypic classes of the QTL model. The specific pairs of alleles at these markers were identified and represent the optimal haplotypes at each QTL, and these can be converted to marker assays for marker-assisted selection of QTL alleles.

Expression analysis using RNA-Seq

The RNA-Seq analyses of two FPR tolerant clones "CATIE-R4" and "CATIE-R6" ("UF 273" Type I × "PA 169") (Phillips-Mora et al. 2013) and two susceptible FPR but tolerant BPR clones "CATIE-1000" and "POUND 7" of unknown pedigrees were part of a previously published study by Ali et al. (2014) using the Belizean Criollo genotype (B97-61/B2) cacao genome assembly (Argout et al. 2011). Data for three independent biological replicates for each of the four clones were included in the analysis. Information related to the *M. roreri* infection assay, RNA isolation, and RNA-Seq analysis can be obtained from the previous study (Ali et al. 2014). RNA reads from RNA-Seq libraries in fastq format were trimmed up using BBDuk version 37.58 (Bushnell 2014), using adapters.fa with parameters ktrrim = r, k = 23, mink = 11, hist = 1, tpe, tbo. Trimmed reads were aligned using HISAT2 2.1.0 (Pertea et al. 2016) to the coding sequences (CDS) of the cacao Matina1-6 genome v1.1 (Motamayor et al. 2013). Tabulated raw counts from each CDS were obtained from the HISAT2 alignment. Raw counts were normalized using the DESeq2 (Anders and Huber 2010) (Galaxy Version 2.11.40.1) available in Galaxy pipeline (https://usegalaxy.org/). Estimation and statistical analysis of the expression level using normalized count data of each gene with three replicates for each library were performed using the DEseq2 package (Supplemental File 2). Genes associated with the QTL were defined as those on either side of the primary marker out to one-half of the distance to the next marker, distinguished by physical position (bp), as determined by the QTL analysis. Gene ontology (GO) analysis was carried out on QTL-associated genes using the program Blast2GO (Conesa et al. 2005). KEGG (Kyoto Encyclopedia of Genes and Genomes) Automatic Annotation Server (KAAS) was used to obtain KEGG Orthology and KEGG pathways involving the genes present in the QTL by BLASTx comparisons against the manually curated KEGG GENES database (Moriya et al. 2007). Identification of pathogen receptor gene orthologs was carried out by BLASTx search against the PRGdb (version 3.0) (Osuna-Cruz et al. 2018).

Results

Phenotypic evaluation of F$_1$ progenies

Mean phenotypic values of frosty pod rot external severity (ES), internal severity (IS), and BPR disease severity (DL10) are presented in Fig. 1. For ES and IS, inoculation data from 169 Type I F$_1$ individuals was used to calculate the mean phenotypic values. In contrast, 175 Type I F$_1$ progenies were employed to estimate BPR DL10 mean phenotypic

values. This difference was due to the low production of pods in these plants as well as the lack of available pods at the time of *M. roreri* inoculations. ES, IS, and DL10 values ranged from 1 (±0.00) to 3.38 (±0.29), 1.05 (±0.05) to 4.83 (±0.13), and 1.02 (±0.01) to 12.84 (±0.79), respectively. The means for ES and IS were 2.06 (±0.48) and 2.81 (±0.86). Also, the number of progenies classified as resistant to highly resistant to FPR (i.e., severity values below 2.0), based on ES and IS severity scores, was 97 and 41, respectively (Fig. 1a, b). Mean DL10 was 3.98 (±2.75) and 110 trees had values that classified them as highly resistant to resistant (Fig. 1c). Most trees were classified as highly resistant to resistant based on ES (97 trees) and DL10 (110 trees).

Polymorphism evaluation and linkage maps

From a final set of 5149 SNPs, 2910 loci were polymorphic between the parental genotypes. One hundred seventy-nine Type I F$_1$ progenies were used to construct the ten-linkage group map. None of the nearest–neighbor stress test (N.N. Stress) values were higher than the threshold value of 2; therefore, no additional SNP loci were eliminated from the map. The number and proportion of segregation types used for the map were the following: 875 (30%) <*lm × ll*>, 664 (23%) <*hk × hk*>, and 1371 (47%) <*nn × np*>. In addition, all linkage groups were composed of SNP markers whose location corresponded to their physical location based on the Matina 1-6 v1.1 genome sequence assembly. Chromosome 5 showed the maximum number of loci (151) segregating in "POUND 7" (<*lm × ll*> maternal segregation type), while chromosome 2 was highest in segregating loci (253) in "UF 273" Type I (<*nn × np*> paternal segregation type). The maximum number of segregation types <*hk × hk*>, indicating that both parents are heterozygous for a locus, was detected on chromosome 8 with 147 loci (Table 1). JoinMap® 5 identified 1883 loci as unique and 1027 loci as identical; however, all loci were included in the final map. The final number of unique loci by chromosome is presented on Table 1. The highest number of unique loci (273) was observed in chromosomes 1 and 2. In contrast, the lowest number (177) was detected in chromosome 3. The total length of the map was 1036.72 cM with a density of 0.55 cM/SNP. Among individual linkage groups, chromosome 1 was the largest with 136.19 cM. Conversely, chromosome 7 was the shortest, with 64.51 cM and 120 non-identical loci. The average chromosome length was 103.67 cM. The average inter-marker distance was 0.93 cM and the largest inter-marker distance of 19.58 cM was observed on chromosome 5 (Table 1, Fig. 2).

Even though all SNP markers mapped to their previously assigned chromosomes based on their physical location on the Matina 1.6 v1.1 assembly scaffolds, a series of changes in marker order were observed across all chromosomes. When comparing the SNPs physical position in the respective

Fig. 1 Distribution of **a**) plant means for FPR external severity, **b**) plant means for FPR internal severity, and **c**) plant means for BPR disease severity among Type 1 F_1 progenies of the cross "POUND 7" × "UF 273"
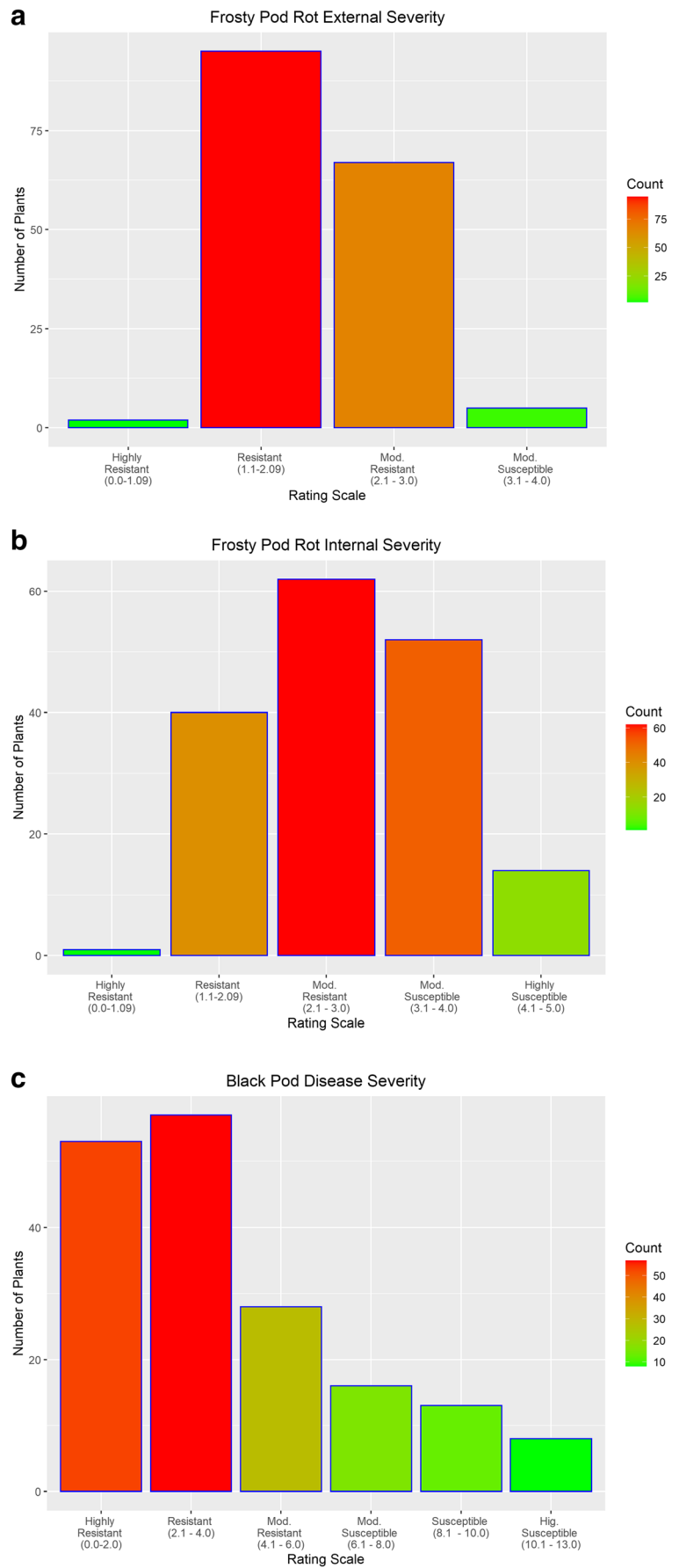
Table 1 Parental and consensus linkage maps of the F$_1$ population of "POUND 7" × "UF 273" Type I constructed with 179 individuals and 2910 SNP loci

Chromosome	"POUND 7" (P$_1$)		"UF 273" Type I (P$_2$)		Segregation pattern			Consensus map					
	No. of loci	Length (cM)	No. of loci	Length (cM)	<lm × ll>	<hk × hk>	<nn × np>	Total No. of loci	No. of unique loci	Length (cM)	Average marker interval (cM)	Gaps ≤ 2 cM	Maximum gap (cM)
1	189*	133.31	332	136.19	105	84	249	438	273	136.19	0.74	176	2.89
2	205	122.72	351	122.72	107	98	253	458	273	122.72	0.75	154	4.02
3	110	114.03	44	72.13	98	15	29	142	110	114.03	1.39	64	7.85
4	169	106.80	225	103.20	89	80	145	314	201	106.80	0.91	107	3.81
5	184	108.63	114	129.50	151	33	81	265	164	129.50	1.26	88	19.58
6	95	82.42	159	88.28	60	35	124	219	144	88.23	0.93	88	5.62
7	92	63.51	124	64.51	36	56	68	160	120	64.51	0.93	64	4.01
8	217	69.02	246	65.704	70	147	103	320	215	69.02	0.55	125	2.00
9	215	132.73	279	126.46	127	88	191	406	266	132.73	0.83	150	4.69
10	59	72.99	156	71.30	32	28	128	188	117	72.99	0.96	69	3.78
Average	153.50	100.62	203	98.00	87.50	66.40	137.10	291	188.30	103.67	0.93	108.50	5.83
Total	1535	1006.16	2030	980.00	875	664	1371	2910	1883	1036.72		1085	

*hk × hk segregation pattern markers are included in both parental maps

Matina 1-6 assembly scaffolds with their corresponding positions in the genetic linkage maps, chromosomes 5 and 6 present the most discrepancies. The marker order for the rest of the chromosomes showed a high degree of collinearity with the Matina 1-6 v1.1 physical map (Supplemental Fig 1, Supplementary Material S1).

SNP markers associated with FPR and BPR resistance

Genome-wide LOD [−Log10(P)] significance thresholds used in the SIM and CIM were calculated by GenStat using a modified Bonferroni correction developed by Li and Ji (2005). Their values at significance levels of $P < 0.05$, $P < 0.01$, and $P < 00.01$ were 3.64, 4.34, and 5.34, respectively for all three traits (ES, IS, and LD10) were studied. SNP markers associated with FPR resistance were first detected in this study with marker regression and SIM using the recommended step size along the genome of 10^6 and 10 cM, respectively (Boer et al. 2017). SNP loci Tcm004s01757744 (chromosome 4, 9.34 cM), Tcm007s10151858 (chromosome 7, 45.98 cM), and Tcm009s41904987 (chromosome 9, 132.73 cM) presented significant associations ($P < 0.01$, $P < 0.01$, and $P < 0.05$) with ES. In addition, SNP markers Tcm002s00690270, Tcm007s10367506, Tcm008s05103982, and Tcm009s41521554 located on chromosomes 2 (2.82 cM), 7, (46.92 cM), 8 (29.67 cM) and 9 (130.20 cM) exhibited significant association with IS ($P < 0.01$, $P < 0.05$, $P < 0.01$, and $P < 0.05$).

In contrast, Tcm004s28852879 on chromosome 4 (85.44 cM) and Tcm010s23394888 on chromosome 10 (65.07 cM) were the only SNP markers identified by marker regression and SIM to be highly significantly associated with BPR DL10 severity ($P < 0.01$ and $P < 0.001$, respectively).

The SNP markers previously identified as associated with FPR and BPR resistance were used as cofactors by GenStat in the CIM analyses. To prevent co-linearity between a cofactor and a potential QTL, a default value of 50 cM was used by GenStat for the minimum cofactor proximity. Results based on the evaluated variables (ES and IS) indicated that eleven QTL associated with resistance to FPR were identified by CIM in this study and are located on chromosomes 2, 4, 7, 8, 9, and 10 (Table 2 and Fig. 3a, b, c, d, e, f). For ES, five QTL were mapped to chromosomes 4, (9.34 cM), 7 (29.10 cM), 8 (31.22 cM), 9 (132.73 cM), and 10 (17.40 cM) with LOD values of at least 3.87 ($P < 0.05$) accounted for 11.19, 10.19, 6.70, 10.23, and 4.96% of the variation, respectively. However, only two SNP markers previously associated with ES by SIM and located on chromosomes 4 (Tcm004s01757744) and 9 (Tcm009s41904987) were detected by CIM as linked with ES (Table 2, Fig. 3b, e).

Six QTL associated with FPR resistance and detected by IS were discovered on chromosomes 2 (17.98 cM), 4 (9.34 cM), 7 (29.10 cM), 8 (31.22 cM), 9 (123.48 cM), and 10 (8.21 cM). These QTL explained 11.08, 8.62, 7.71, 10.54, 9.12, and 5.54% of the variation with significant LOD values of at least 4.50 ($P < 0.01$), respectively. Moreover, SNP markers Tcm004s01757744, Tcm007s05825365, and
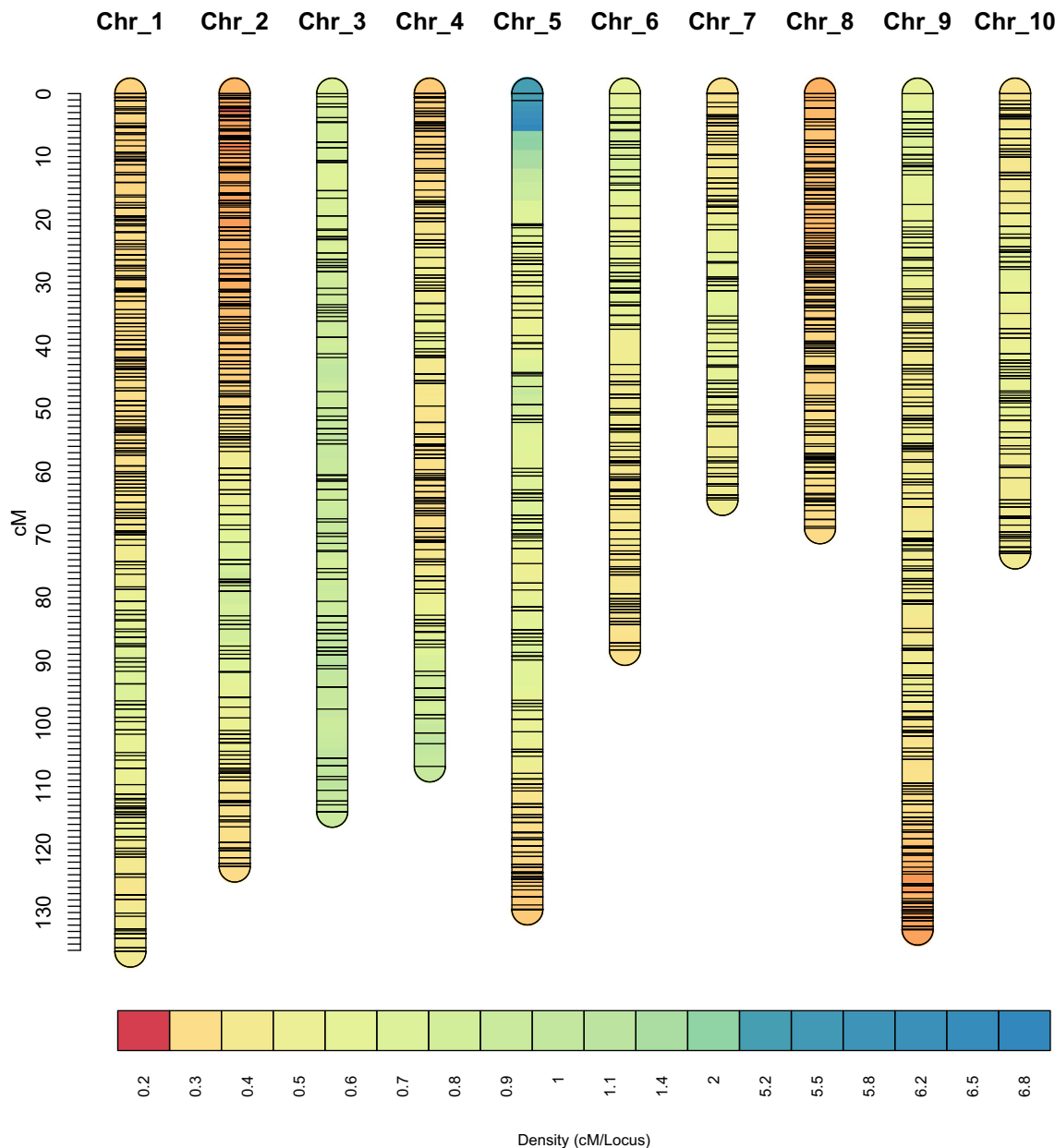
Fig. 2 Genetic linkage map constructed using 179 Type 1 F$_1$ progenies derived from the cross: "POUND 7" × "UF 273"

Tcm008s05380821, associated with ES, were also linked to IS and used as cofactors in the CIM analysis for both traits. On chromosomes 9 and 10, SNP markers associated with both ES and IS traits were found at different locations. Only chromosome 2, presented SNP markers associated with IS resistance (Table 2 and Fig. 3a, b, c, d, e, f).

Four QTL for BPR resistance, based on the DL10 variable, were identified on chromosomes 2, 4, 8, and 10 using CIM (Table 2 and Fig. 3a, b, d, f). SNP locus Tcm002s08313597 located at 43.60 cM on chromosome 2 is linked to this first QTL ($P < 0.05$ and 8.31%). SNP locus Tcm004s28538741 ($P < 0.01$ and 11.01%) is located on the second QTL, at 78.67 cM on chromosome 4. The third QTL ($P < 0.001$ and

8.23%) is located on Tcm008s04656460 at 27.12 cM. The last QTL (Tcm010s22418501) was detected on chromosome 10 at 61.02 cM with an LOD value of 10.88 ($P < 0.01$) and accounted for 23.07% of the variation. All these SNP markers associated with the previously described QTL were used as cofactors on the CIM analyses.

In general, significant additive effects due to the average difference between the two alleles (a, b) inherited from "POUND 7" ($a_1 = gp\_additive$) and the other two alleles (c, d) from "UF 273" Type I ($a_2 = gp\_additive2$) were detected for each of the QTL associated with ES, IS, and DL10. The largest additive effect was observed on chromosome 10 (1.18) and was associated with BPR DL10 ("POUND 7"). Four QTL

Table 2 Quantitative trait loci calculated with CIM mapping for FPR external severity, FPR internal severity, and BPR disease severity in an F_1 population of "Pound 7" × "UF 273" Type I

Chr.	Trait	Map position of the QTL	Peak marker	−Log10(P) score	Additive effect Pound 7 a_1	Additive effect UF 273 a_2	Dominance effects d	Genotypic mean ac	Genotypic mean ad	Genotypic mean bc	Genotypic mean bd	% var. expl.
4	FPR ES-1	9.34	Tcm004s01757744	5.11**	0.06	−0.11	−0.08	2.04	2.00	2.34	1.96	11.19
7	FPR ES-2	29.10	Tcm007s05825365	7.06***	−0.03	−0.14	0.09	2.35	1.89	2.09	2.01	10.19
8	FPR ES-3	31.22	Tcm008s05380821	4.53**	0.12	−0.01	−0.08	1.89	2.03	2.29	2.12	6.70
9	FPR ES-4	132.73	Tcm009s41904987	7.42***	0.02	−0.17	—	2.23	1.89	2.28	1.93	10.23
10	FPR ES-5	17.40	Tcm010s02288537	3.87*	0.06	0.08	—	1.90	2.06	2.10	2.27	4.96
2	FPR IS-1	17.98	Tcm002s04088162	7.57***	0.06	−0.27	—	3.25	2.70	3.10	2.55	11.08
4	FPR IS-2	9.34	Tcm004s01757744	4.97**	0.10	−0.20	—	2.99	2.60	3.20	2.81	8.62
7	FPR IS-3	29.10	Tcm007s05825365	6.14***	−0.06	−0.23	0.12	3.30	2.62	2.96	2.73	7.71
8	FPR IS-4	31.22	Tcm008s05380821	9.19***	0.29	−0.10	−0.18	2.52	2.70	3.47	2.91	10.54
9	FPR IS-5	123.48	Tcm009s40433054	8.31***	0.05	−0.26	0.14	3.25	2.44	3.08	2.83	9.12
10	FPR IS-6	8.21	Tcm010s00936105	4.50**	0.20	0.11	—	2.59	2.81	3.00	3.21	5.54
2	BPR DL10-1	43.60	Tcm002s08313597	3.82*	0.22	−0.09	−0.65	3.31	4.43	5.05	3.57	8.31
4	BPR DL10-2	78.67	Tcm004s28538741	4.50**	−0.23	0.73	−0.35	3.24	5.39	3.48	4.25	11.01
8	BPR DL10-3	27.12	Tcm008s04656460	6.14***	0.42	0.78	—	2.90	4.45	3.73	5.28	8.23
10	BPR DL10-4	61.02	Tcm010s22418501	10.88****	1.18	−0.05	—	2.96	2.86	5.32	5.22	23.07

*, **, ***, **** Significance levels at the $P < 0.05$, $P < 0.01$, and $P < 0.001$ were 3.64, 4.34, and 5.34, respectively
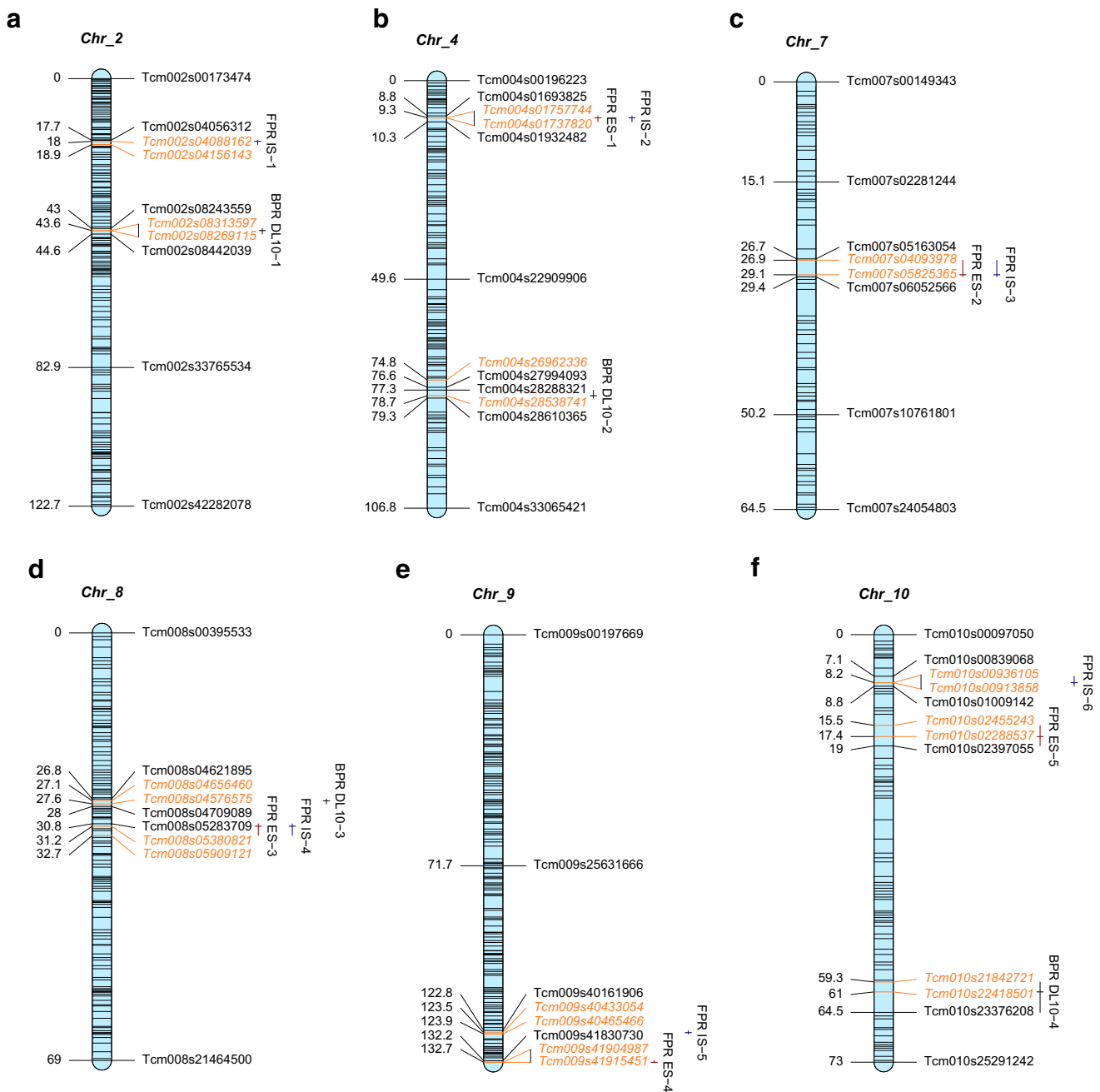
Fig. 3 Locations of QTL associated with FPR external severity (ES), FPR internal severity (IS), and BPR disease severity (DL10) in a Type 1 F_1 progenies of the cross "POUND 7" × "UF 273." The SNPs marked with colors are the peak marker and flanking marker of each QTL

associated with ES presented minor dominance effects. In contrast, 3 IS QTL as well as 2 DL10 QTL showed bigger dominance effects ($d = gp\_dominance$) which is due to the within-locus allelic interaction effect. The largest dominance effects were observed on chromosomes 2 (−0.65) and 4 (−0.35) and associated with two QTL for DL10 (Table 2).

A segregation ratio of 1:1 was observed on majority the SNP markers associated with QTL for ES and IS except for FPR IS-6. On the other hand, a 1:2:1 segregation ratio was presented in three of the four SNPs associated with QTL for DL10.

Each QTL peak is defined by a single SNP marker with two alleles, with the pair of four parental haplotypes inferred at this position through the linkage analysis, leveraging the information from flanking markers. Therefore, we identified a pair of SNPs in the region of each QTL peak that jointly can discriminate the four parental haplotype combinations, and the specific two-locus genotype that uniquely identifies the most resistant genotype at each QTL (Table 3). Moving from single SNPs to SNP pairs at the QTL peaks allows us to design SNP assays that permit marker-assisted selection of the favored

genotypes at each QTL. Six of the selected SNP marker pairs are completely linked, whereas the remaining nine SNP pairs are separated by 0.39 to 3.84 cM (Table 3).

QTL gene composition and expression

A summary of the QTL compositions is presented in Supplementary Tables 1 and 2 including gene identifications based on GO and KEGG analysis and gene expression profiles (Supplemental Material S3). The gene numbers of the various QTL varied widely ranging from 11 to 185. There were four chromosomal rearrangements involving markers associated with specific QTL which complicated identification of associated genes: QTL DL10-1, DL10-3, ES-5, and ES-2/IS-3. The chromosomal rearrangements were all relatively close to the primary QTL marker based on the cacao Matina1-6 genome v1.1. Most of the genes within the QTL were expressed with at least one read. There were a limited number of genes in each QTL showing consistent differential expression between resistant ("CATIE-R4" and "CATIE-R7") and susceptible ("POUND 7" and "CATIE-1000") clones (Supplemental Material S2 and S3).

Among the various gene groupings, associated with the GO analysis of biological function frequently identified (Supplemental Tables 1 and 2), oxidation/reduction processes and proteolysis stand out as relevant to disease resistance. Many integral membrane proteins were identified based on the GO cellular component. Genes with putative kinase activity stand out among the classes identified by the GO molecular function analysis. Every QTL had at least 1 gene identified (e-value ≤ -0.50, identity $\geq 40\%$) as a potential disease-resistant associated gene in the pathogen receptor database PRGdb, with 40 genes being identified in ES-2/IS-3 and 28 being identified in DL10-4. KEGG IDs were identified for 143 of the QTL-associated genes.

Many of the genes located within the QTL can easily be associated with plant defense functions, forming gene clusters on specific QTL (Supplemental Tables 1 and 2): chitinases (IS-5), cysteine/histidine-rich c1 domain (ES-5), cytochrome P450 enzymes (DL10-3), indole-3-acetic acid-amido synthetase (ES-2/IS-3), malectin/receptor protein kinase (IS-6), AAA-ATPase (IS-1), serine carboxypeptidase 44 (ES-2/IS-3), and especially disease resistance/LRR containing proteins (ES-5, ES-2/IS-3, DL10-2, DL10-4). In addition, there were many genes potentially associated with disease resistance which occurred in related sets of two or three on individual QTL (Supplemental Tables 1 and 2). Of interest are the NRT1/PTR family 5.4 (DL10-4), plant self-incompatibility S1 (IS-1), and TPD1 protein homolog 1 genes (DL10-2) with members showing downregulation in resistant clones and non-specific lipid-transfer proteins (ES-4) with members showing upregulation in resistant clones.

Discussion

A consensus linkage map was built from an F_1 population of 179 individuals from a cross between "UF 273" Type I and "POUND 7" containing 1883 unique loci and spanning 1036.72 cM, with an average marker interval of 0.93 cM. Using the same set of SNPs utilized in the present study, Royaert et al. (2016) constructed a linkage map with 459 individuals of an F_1 segregating population of "TSH 1188" × "CCN 51" and obtained a smaller map of 852.8 cM with 3526 loci. Since the number of polymorphic markers (68%) between the population's parents in Royaert et al. (2016) study was higher, the final map's size could have been influenced by this condition. In contrast, a linkage map (1268 cM) with progeny from the "POUND 7" × "UF 273" Type II cross was constructed by Livingstone et al. (2015). However, the map presented here and obtained with the Type I progeny is smaller in size (1037 cM). The difference in size between these two maps could be due to the number of individuals (179 of Type I versus 68 of Type II) used to construct them. Also differences observed between progeny linkage maps from "UF 273" Type I and Type II parents could be due to the number of SNP loci (346) that segregate uniquely in Type I progeny versus those that segregate only in Type II progeny (260). Except for LG5, all the linkage groups on the "POUND 7" × "UF 273" Type II map were larger than the one assembled in this study with the "UF 273" Type I parent progeny. In the majority of linkage groups of the Type I population, the maximum gap length was also smaller, except for LG5 and LG6 (Livingstone et al. 2015). In addition, Livingstone et al. (2015) and Royaert et al. (2016) removed SNP (7 and 17) loci from their linkage maps due to nearest–neighbor stress test (N.N. Stress) values being higher than the threshold value of 2. In the map presented here, no nearest–neighbor stress test value higher than 2 was observed, indicating a map of good quality. Lastly, due to the fact that "UF 273" Type II parent was classified as moderately resistant (Phillips-Mora et al. 2013), the Type II progeny utilized by Livingstone et al. (2015) could not be used to find QTL associated with FPR and BPR resistance.

Recently, Livingstone et al. (2017) using a 15K SNP Illumina chip screened two F_1 populations, "EET 95" × "Silecia 1" and "SCA 12" × "Unknown," composed of 576 and 238 individuals, respectively. The first one generated a 3636 SNP loci map of 834 cM length; the second one produced a 1269-cM map with 1862 SNP loci. The maximum gap length for all linkage groups presented in the current study is smaller than the ones in the "SCA 12" × "Unknown" F_1 population with the exception of LG3 and LG5 (Livingstone et al. 2017). However, the maximum gap lengths of the "EET 95" × "Silecia 1" population linkage groups are smaller for all of the linkage groups when compared to the linkage group maps presented here, except for LG8 (Livingstone et al.

Table 3 Two-locus genotype combinations for the parents "Pound 7," "UF 273" Type I and mean values for four offspring genotypes defined by pairs of SNP markers associated with QTL for FPR external severity, FPR internal severity, and BPR disease severity. Informative marker pairs were chosen to distinguish the four possible haplotypic combinations at each QTL

Chr.	QTL	cM position	Marker pair defining QTL genotypes	Segregation pattern	Pound 7 genotype	UF 273 T1 genotype	Genotype *ac* (mean)	Genotype *ad* (mean)	Genotype *bc* (mean)	Genotype *bd* (mean)
4	FPR ES-1	9.34	Tcm004s01757744	<nn × np> (1:1)	TT	TC	TT (1.98)	TC (1.92)	TT (2.35)	TC (1.97)
		9.34	Tcm004s01737820	<lm × ll> (1:1)	GA	AA	GA	GA	AA	AA
7	FPR ES-2	29.10	Tcm007s05825365	<nn × np> (1:1)	CC	CT	CC (2.29)	CT (1.86)	CC (2.12)	CT (1.96)
		26.91	Tcm007s04093978	<lm × ll> (1:1)	AC	CC	AC	AC	CC	CC
8	FPR ES-3	31.22	Tcm008s05380821	<lm × ll> (1:1)	AG	AA	AA (1.85)	AA (2.01)	GA (2.21)	GA (2.17)
		32.67	Tcm008s05909121	<nn × np> (1:1)	CC	TC	CT	CC	CT	CC
9	FPR ES-4	132.73	Tcm009s41904987	<nn × np> (1:1)	CC	CT	CC (2.24)	CT (1.94)	CC (2.22)	CT (1.88)
		132.73	Tcm009s41915451	<lm × ll> (1:1)	AG	GG	AG	AG	GG	GG
10	FPR ES-5	17.40	Tcm010s02288537	<nn × np> (1:1)	GG	GA	GG (1.88)	GA (2.09)	GG (2.06)	GA (2.23)
		15.52	Tcm010s02455243	<lm × ll> (1:1)	AG	AA	AA	AA	GA	GA
2	FPR IS-1	17.98	Tcm002s04088162	<nn × np> (1:1)	AA	AG	AA (3.19)	AG (2.37)	AA (3.03)	AG (2.63)
		18.93	Tcm002s04156143	<lm × ll> (1:1)	GA	GG	GG	GG	AG	AG
4	FPR IS-2	9.34	Tcm004s01757744	<nn × np> (1:1)	TT	TC	TT (2.73)	TC (2.56)	TT (3.26)	TC (2.68)
		9.34	Tcm004s01737820	<lm × ll> (1:1)	GA	AA	GA	GA	AA	AA
7	FPR IS-3	29.10	Tcm007s05825365	<nn × np> (1:1)	CC	CT	CC (3.13)	CT (2.48)	CC (3.01)	CT (2.65)
		26.91	Tcm007s04093978	<lm × ll> (1:1)	AC	CC	AC	AC	CC	CC
8	FPR IS-4	31.22	Tcm008s05380821	<lm × ll> (1:1)	AG	AA	AA (2.42)	AA (2.67)	GA (3.16)	GA (3.00)
		32.67	Tcm008s05909121	<nn × np> (1:1)	CC	TC	CT	CC	CT	CC
9	FPR IS-5	123.48	Tcm009s40433054	<lm × ll> (1:1)	TC	CC	TC (3.20)	TC (2.50)	CC (2.97)	CC (2.70)
		123.87	Tcm009s40465466	<nn × np> (1:1)	TT	CT	TC	TT	TC	TT
10	FPR IS-6	8.21	Tcm010s00936105	<hk × hk> (1:2:1)	AG	AG	AA (2.46)	AG (2.85)	GA (2.85)	GG (3.14)
		8.21	Tcm010s00913858	<nn × np> (1:1)	CC	CT	CC	CT	CC	CT
2	BPR DL10-1	43.60	Tcm002s08313597	<hk × hk> (1:2:1)	CT	CT	CC (3.39)	CT (4.51)	TC (5.12)	TT (3.11)
		43.60	Tcm002s08269115	<nn × np> (1:1)	CC	CT	CC	CT	CC	CT
4	BPR DL10-2	78.67	Tcm004s28538741	<lm × ll> (1:1)	TC	TT	TT (3.15)	TT (5.59)	TC (3.23)	TT (4.03)
		74.83	Tcm004s26962336	<nn × np> (1:1)	CC	TC	CT	CC	CT	CC
8	BPR DL10-3	27.12	Tcm008s04656460	<hk × hk> (1:2:1)	CT	CT	CC (2.75)	CT (4.42)	TC (3.68)	TT (5.09)
		27.55	Tcm008s04576575	<lm × ll> (1:1)	TC	CC	TC	TC	CC	CC
10	BPR DL10-4	61.02	Tcm010s22418501	<hk × hk> (1:2:1)	AG	AG	AA (2.51)	AG (2.74)	GA (5.34)	GG (5.21)
		59.31	Tcm010s21842721	<nn × np> (1:1)	AA	GA	AG	AA	AG	AA

2017). Finally, the results presented here confirmed the findings of Royaert et al. (2016) in regards to chromosomes size, where the smallest ones were 6, 7, 8, and 10. Because the linkage map reported here was constructed with the recent release of JoinMap® 5, it provides a more accurate and efficient linkage map construction since a very fast deterministic EM algorithm for the multipoint recombination frequency estimation in the ML mapping procedure has been implemented in this new version (Van Ooijen 2018).

Earlier studies examining the inheritance of genes conferring frosty pod rot resistance suggested that it could be recessive and/or polygenic in nature, since most of the progenies were classified as susceptible. However, these studies were conducted using a small number of segregating populations and accessions (Phillips-Mora and Castillo 1999). In contrast, Phillips-Mora et al. (2017) using larger and more diverse germplasm collection stated that the resistance to FPR was polygenic and mainly additive. In the current study, with a near-normal distribution of the FPR IS values and a low skewness value (0.26), forty-one (24%) individuals were classified as highly resistant or resistant (based on internal symptom development, Fig. 1b), indicating that this trait could be controlled by a few genes or by environmental effects which may have affected the artificial inoculation process. In the current study, the size of the additive effects from "UF 273" Type I indicates that it contributed significantly to most of the QTLs associated with FPR resistance.

In contrast, the distribution of BPR resistance (Fig. 1c), based on lesion diameter 10 days after inoculation, is biased toward the resistant parent (skewness value 1.17) because 110 (57%) of the F_1 progenies were categorized as highly resistant or resistant. Resistance to BPR has been reported in many cacao accessions worldwide (Pound 1936; Spence 1961). Qualitative and quantitative inheritance both play a part in the disease resistance for BPR (Enríquez and Soria 1984). Tan and Tan (1990) indicated that the resistance to BPR was inherited in a quantitative manner. Iwaro and Butler (2000), in a subsequent study, concluded that BPR resistance was inherited as a quantitative trait, with narrow sense and broad sense heritability estimated to be 0.33 and 0.51, respectively. In addition, there is evidence that both additive and non-additive gene action are involved in the inheritance of BPR resistance. Transgressive segregation has been observed in several crosses ("SCA 6" × "CATONGO," "POUND 7" × "UF 676," "POUND 7" × "CATONGO," and "UF 613" × "UF 676"), with progeny showing greater resistance than either parent (Enríquez and Soria 1999). In the present study, additive effects from "POUND 7" contributed significantly only to three QTLs for FPR resistance located in chromosomes 4 and 8. Dominance effects were larger in three QTLs for resistance to FPR (FPR-IS-3, 4, and 5) and two for resistance to BPR (BPR DL10-1 and 2).

Transgressive segregation evidence for BPR has also been observed here since 53 of the F_1 progenies were categorized as highly resistant.

Information regarding the pedigree of "POUND 7" does not exist since it was collected in the Peruvian Amazon (Pound 1943). In contrast, UF 273 was developed by the United Fruit cacao breeding program in Costa Rica; however, no pedigree information is available (Johnson et al. 2007). Recent reports indicate the clear genetic differences between the two parents in the current study. According to Mata-Quirós et al. (2018), the genetic background of "UF 273" Type I consists of 63% of alleles from the Nacional and 32% from the Amelonado genetic groups, respectively. In contrast, "POUND 7" alleles are approximately 2/3 from the Nanay and 1/3 from Iquitos genetic groups (Cornejo et al. 2018). Cacao clones belonging to other genetic groups has been screened for FPR resistance. Phillips-Mora et al. (2017) reported that fifty clones categorized as resistant and moderately resistant to FPR were classified as Amelonado, Criollo, Curaray, Guiana, Iquitos, Marañon, Nacional, and Purús or as an admixture of these genetic groups. However, Romero Navarro et al. (2017) conducted a genome-wide association study (GWAS) using 148 cacao clones with mainly Amelonado and National genetic backgrounds and stated that only the Nacional background was significantly correlated with field resistance to FPR but also associated with BPR susceptibility. According to Phillips-Mora et al. (2017), the steady appearance of cacao clones resistant and moderately resistant to FPR in areas where the disease is not present could indicate that non-specific R genes are mainly responsible for FPR. A proposed *M. roreri* center of diversity area (Phillips-Mora 2003) has been already considered an important region for germplasm collection expedition in the search of FPR resistant accessions. However, none of the fifty cacao accessions resistant and moderately resistant to FPR came from the *M. roreri* center of origin, located in Eastern Colombia (Phillips-Mora et al. 2017). Currently, "CATIE-R6" is the only available clone that is resistant to both FPR and BPR. In contrast, "CATIE-R4" is resistant to FPR but susceptible to BPR. Their pedigree is "UF 273" Type I × "PA 169" (Phillips-Mora et al. 2013). Romero Navarro et al. (2017) also observed that "UF 273" Type I was in the pedigree of 60% of the cacao highest-yielding clones of the CATIE breeding program. "CATIE-R73" ("PA 169" × "ARF 22"), a highly resistant clone to BPR, was also evaluated in this study. One of its progenitors, "ARF 22," is a BPR resistant accession and the product of a cross of "UF 613" × "POUND 7."

Phillips-Mora et al. (2017) reported the clone "ARF 33" ("POUND 7" × "SCA 6") as one of the fifty clones characterized as resistant and moderately resistant to FPR. Since "POUND 7" is one of the parents, it could indicate that "POUND 7" is also contributing to the resistance to FPR.

Based on the previously cited results, there is enough evidence to confirm the contribution of "UF 273" Type I and "POUND 7" accessions as one of the primary sources of FPR and BPR resistance genes used in current cacao breeding programs in the Americas.

Brown et al. (2007) genotyped 256 F_1 progenies in the population "POUND 7" × "UF 273" (Types I and II) with 180 SSR markers. They detected a total of eight QTL for FPR and BPR resistance. In the current study, 15 QTL for resistance to FPR and BPR were identified on multiple chromosomes. However, only the QTL for FPR ES mapped to chromosome 2, as reported by Brown et al. (2007), and it was not validated. The remaining three QTLs located on chromosomes 7 and 8 for FPR ES, as well as on chromosomes 2 and 8 for FPR IS, were validated. Additionally, seven previously unreported QTL for FPR resistance were observed here: two on chromosome 4 (FPR ES-1 and IS-2), one on chromosome 7 (FPR IS-3), two on chromosome 9 (FPR ES-4 and FPR IS-5), and two on chromosome 10 (FPR ES-5 and FPR IS-6). Since the paternal parent ("UF 273" Type I) linkage map (<$nn × np$>) contains more loci (1371 vs. 875) than the maternal parent ("POUND 7") linkage map (<$lm × ll$>), it is possible this may have contributed to the finding of more QTL associated with "UF 273" Type I. Interestingly, only one SNP associated with QTL for ES and IS had a <$hk × hk$> segregating ratio. In contrast, most of the SNP associated with QTL for DL10 were of <$hk × hk$> segregating ratio. A scale of 1 to 5 was used to evaluate ES and IS; however, DL10 was measured quantitatively (cm), this difference in trait measurement methodology could be responsible for the high number of BPR linked loci with <$hk × hk$> segregating ratio since traits measured quantitatively could help to differentiate more easily the segregating genotypes present in the evaluated population.

Romero Navarro et al. (2017) using GWAS methodologies reported that six SNP loci were significantly linked with FPR under natural field disease incidence and they were located on chromosomes 4, 5, 6, and 10. Later, McElroy et al. (2018) in a genome-wide association and genomic selection study, composed of three related populations, reported that SNP markers associated with FPR were located on chromosomes 1, 3, 5, 9, and 10. In the current study, only two QTL associated with FPR were located on chromosomes 4 and 10 but they were located in different regions of the chromosomes as the one reported by Romero Navarro et al. (2017). When comparing with the results of McElroy et al. (2018), only two QTL for FPR in this study mapped to chromosomes 9 and 10. In both studies, this is possibly due to the different loci segregating in the experiments. In addition, Romero Navarro et al. (2017) and McElroy et al. (2018) used different germplasms, with different genes of resistance than those compared here. Another source of study variation may derive from the use of artificial inoculations versus the natural field disease

incidence as used in the study of Romero Navarro et al. (2017) as well as in McElroy et al. (2018). For example, Romero Navarro et al. (2017) noted that the incidence of FPR in "UF 273" Type I progeny under natural infection is very low, limiting the ability to recognize differences (McElroy et al. 2018).

FPR resistance QTL on chromosomes 4, 7, and 8 (FPR ES-1, FPR IS-2, FPR IS-3, FPR ES-3, and FPR IS-4, Fig. 3b, c, and d) were identified using both measures of disease severity (internal vs. external) and shared the same SNP loci (Tcm004s01757744, Tcm007s05825365, and Tcm008s05380821), indicating a large degree of overlap, if not identity, between the genes contained in each. In contrast, the SNPs associated with ES and IS in chromosomes 9 and 10 were different. A strong correlation between internal and external severity lesions was reported by Phillips-Mora (1986), with internal symptom development being slightly more responsive to artificial inoculation. It may be better to select resistant germplasm based on the FPR IS QTL, since this presented a higher LOD score as well as larger additive effects. The internal nature of FPR IS allows the bean to be reached easily.

Results presented here further support the importance of using quantitative measurements during the evaluation of FPR. Comparisons of the qualitative (external severity) versus quantitative (internal severity) measurements clearly indicated that the latter are more accurate for the evaluation of resistance to FPR. A total of one hundred and sixty-three trees used in determining FPR external severity ranked as moderately resistant or higher whereas one hundred and four of these trees ranked as moderately resistant or greater based on FPR internal severity.

Great efforts have been placed on identifying QTL for resistance to black pod rot (Risterucci et al. 2000; Brown et al. 2007; Lanaud et al. 2009; Akaza et al. 2016; Motilal et al. 2016). Brown et al. (2007) previously reported QTL for BPR on chromosomes 4, 8, and 10. In the present study, four QTL associated with BPR DL10 were identified on chromosomes 2, 4, 8, and 10. A QTL (BPR DL10-1) on chromosome 2 was the only one not previously reported by Brown et al. (2007). Significant additive effects due to "POUND 7" were observed in all QTL for BPR resistance. BPR DL10-4 (chromosome 10) was associated with the largest additive effects (1.18) associated with "POUND 7" and accounted for 23.07% of the variation. Also, on BPR DL10-2 and BPR DL10-3 additive effects from "UF 273" Type I contributed to the BPR resistance and dominance effects were present on QTL BPR DL10-1 and BPR DL10-2. Additionally, Akaza et al. (2016), using "POUND 7" as one of the parents ("POUND 7" × "ICS 95"), found a QTL on chromosome 10 for BPR resistance based on results from leaf disk inoculations. They also found QTL in chromosome 4 based on field incidence for the two crosses involving "POUND 7" ("POUND 7" × "ICS 100" and "POUND 7" × "ICS 95"). Recently, Barreto et al. (2018) also using a leaf disk inoculation methodology in an F_1

population of "TSH 1188" × "CCN 51," screened with *Phytophthora palmivora*, *Phytophthora citrophthora*, and *Phytophthora capsici*, and detected six QTL for BPR resistance; however, only two of these were located on chromosomes 2 and 4 and they were both associated with resistance to *P. capsici*.

In the current study, five QTL for FPR ES, the six QTL for FPR IS, and the four QTL for BPR DL10 accounted for 43.27%, 52.61%, and 50.62% of the phenotypic variation, respectively. The SNP marker pairs associated with these traits could be used in a genomic-assisted breeding program to select for ES, IS, and DL10. They also could be validated on clones recently developed by CATIE that used "POUND 7" and "UF 273" Type I as parents and have presented good level of FPR and BPR resistance. In addition, the percentage of the variance accounted for by these markers in the current study is greater than the that reported by Brown et al. (2007) who used fewer molecular markers (180 vs. 2910 loci) but a larger population size (256 vs. 179). Even though the population size used was smaller when compared with previously described maps, the linkage map presented here was adequate for the conduction of this QTL study based on the comparison with previously published maps. However, one of the consequences of a smaller population size is the reduction of the power of detection. Due to this situation, probably fewer QTL were detected in this study even though the number of loci was sixteen times higher. Furthermore, the effect of the observed QTL is likely to be overestimated. This situation is also known as the Beavis effect (Beavis 1994). An example of this Beavis effect may be the cause of the QTL for FPR ES reported by Brown et al. (2007) and located on chromosome 2. This QTL was the only one not validated in the current study. A smaller population size (179 vs. 256) may be responsible for this outcome.

Although it is hard to rule out any gene for potential impact on plant defense, several well-established disease resistance-associated genes, often forming gene clusters or islands, were identified as candidates in this study. The read alignments are associated with the Matina 1-6 v1.1 genome (Motamayor et al. 2013) and unique genes associated with specific clones would naturally be missing from the alignments. This seems to be a likely possibility considering the gene clusters identified, the presence of retroelements, and the observed rearrangements. Although the genes observed to be differentially expressed should be considered, it is likely that the most important aspect of the RNA-Seq analysis was the verification of expressed genes and their potential roles in disease resistance.

In conclusion, the QTL and SNP associated marker pairs reported here represent an important resource for use in genomic-assisted breeding, facilitating selection of cacao resistant to frosty pod rot and black pod rot.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11295-021-01507-w.

Declarations

Conflict of interest The authors declare no competing interests.

References

Akaza MJ, N'Goran JAK, N'Guetta SPA, Kebe BI, Tahi GM, Sangaré A (2009) Resistance to Phytophthora palmivora (Butler) Butler assessed on leaf discs of cacao (Theobroma cacao L.) hybrids trees. Asian Journal of Plant Pathology 3:106–118

Akaza JM, Kouassi AB, Akaffou DS, Fouet O, N'Guetta AS-P, Lanaud C (2016) Mapping QTLs for black pod (Phytophthora palmivora)

resistance in three hybrid progenies of cocoa (Theobroma cacao L.) using SSR markers. Int J Sci Res Publ 6:298–311

Ali SS, Melnick RL, Crozier J, Phillips-Mora W, Strem MD, Shao J, Zhang D, Sicher R, Meinhardt L, Bailey BA (2014) Successful pod infections by Moniliophthora roreri result in differential Theobroma cacao gene expression depending on the clone's level of tolerance. Mol Plant Pathol 15:698–710. https://doi.org/10.1111/mpp.12126

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11:R106. https://doi.org/10.1186/gb-2010-11-10-r106

Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova SN, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto JF, Sabot F, Kudrna D, Ammiraju JSS, Schuster SC, Carlson JE, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Berard A, Viot C, Boccara M, Risterucci AM, Guignon V, Sabau X, Axtell MJ, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza JM, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie WR, Guiderdoni E, Quetier F, Panaud O, Wincker P, Bocs S, Lanaud C (2011) The genome of Theobroma cacao. Nat Genet 43:101–108 http://www.nature.com/ng/journal/v43/n2/abs/ng.736.html#supplementary-information

Bailey BA, Evans HC, Phillips-Mora W, Ali SS, Meinhardt LW (2018) Moniliophthora roreri, causal agent of cacao frosty pod rot. Mol Plant Pathol 19:1580–1594. https://doi.org/10.1111/mpp.12648

Barreto MA, Rosa JRBF, Holanda ISA, Cardoso-Silva CB, Vildoso CIA, Ahnert D, Souza MM, Corrêa RX, Royaert S, Marelli J, Santos ESL, Luz EDMN, Garcia AAF, Souza AP (2018) QTL mapping and identification of corresponding genomic regions for black pod disease resistance to three Phytophthora species in Theobroma cacao L. Euphytica 214:188. https://doi.org/10.1007/s10681-018-2273-5

Bartley BGD (2005) The genetic diversity of cacao and its utilization. CABI, Wallingford. https://doi.org/10.1079/9780851996196.0000

Beavis W (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Proceedings of the forty-ninth annual corn & sorghum industry research conference. American Seed Trade Association, Washington, DC, pp 250–266

Bejarano G (1961) Métodos de inoculación artificial y factores favorables para la infección de Monilia roreri Cif y Par. Universidad Central del Ecuador

Boer M, Cave V, Jansen H, Malosetti M, Mathews K, Murray D, van Eeuwijk F, Welham S (2017) A guide to QTL analysis in Genstat®, 19th edn. VSN International, Hemel Hempstead, UK

Brown JS, Phillips-Mora W, Power EJ, Krol C, Cervantes-Martinez C, Motamayor JC, Schnell RJ (2007) Mapping QTLs for resistance to frosty pod and black pod diseases and horticultural traits in Theobroma cacao L. Crop Sci 47:1851–1858. https://doi.org/10.2135/cropsci2006.11.0753

Bushnell B (2014) BBMap: a fast, accurate, splice-aware aligner. Ernest Orlando Lawrence Berkeley National Laboratory. CA (US), Berkeley

Campbell AS, Ploetz R, Gutierrez OA (2015) Using digital image analysis to evaluate disease resistance. Paper presented at the The Americas Cacao Breeders' Working Group 2nd Annual Meeting, San Salvador, El Salvador, September 9 – 11, 2015

Cervantes-Martinez C, Brown JS, Schnell RJ, Phillips-Mora W, Takrama JF, Motamayor JC (2006) Combining ability for disease resistance, yield, and horticultural traits of cacao (Theobroma cacao L.) clones. J Am Soc Hort Sci 131:231–241

Cheesman E (1944) Notes on the nomenclature, classification and possible relationships of cocoa populations. Tropical Agriculture, Trinidad and Tobago 21:144–159

Coe SD, Coe MD (2013) The true history of chocolate. Thames & Hudson

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Cope FW (1984) Cacao Theobroma cacao (Sterculiaceae). In: Simmonds NW (ed) Evolution of crop plants. Longman, London, UK, pp 285–289

Cornejo OE, Yee M-C, Dominguez V, Andrews M, Sockell A, Strandberg E, Livingstone D, Stack C, Romero A, Umaharan P, Royaert S, Tawari NR, Ng P, Gutierrez O, Phillips W, Mockaitis K, Bustamante CD, Motamayor JC (2018) Population genomic analyses of the chocolate tree, Theobroma cacao L., provide insights into its domestication process. Communications Biology 1:167. https://doi.org/10.1038/s42003-018-0168-6

Crouzillat D, Menard B, Mora A, Phillips W, Petiard V (2000) Quantitative trait analysis in Theobroma cacao using molecular markers. Euphytica 114:13–23

Enríquez GA, Soria J (1984) Mejoramiento genético para resistencia a cinco enfermedades del cacao. Revisión de literatura vol 9. Materiales de Enseñanza No. 9. CATIE, Turrialba, Costa Rica

Enríquez G, Soria J (1999) Genetic research on cocoa diseases at CATIE (1960–1990). In: Proceedings international workshop on the contribution of disease resistance to cocoa variety improvement, Salvador, BA, Brasil, 24–26 Nov 1996, vol 633.7445 I61 1996. INGENIC, pp 33–40

Evans HC (2007) Cacao diseases-the trilogy revisited. Phytopathology 97:1640–1643

Flament MH, Kebe I, ClTment D, Pieretti I, Risterucci AM, N'Goran JAK, Cilas C, Despreaux D, Lanaud C (2001) Genetic mapping of resistance factors to Phytophthora palmivora in cocoa. Genome 44: 79–85. https://doi.org/10.1139/gen-44-1-79

Galindo JJ, Enriquez GA (1984) Field resistance of cacao clones to Monilia roreri Cif & Par in Costa Rica. In: APS (ed) The American Phytopathological Society. Caribbean Division. Annual meeting, San José, Costa Rica, October 22-26, 1984. APS

Gutiérrez OA, Campbell AS, Phillips-Mora W (2016) Breeding for disease resistance in cacao. In: Bailey BA, Meinhardt LW (eds) Cacao diseases: a history of old enemies and new encounters. Springer International Publishing, Cham, pp 567–609. https://doi.org/10.1007/978-3-319-24789-2_18

Haldane J (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. J Genet 8:299–309

Holliday P (1954) The control of black pod and witches' broom disease in Trinidad. In: Reunion del Comite Tecnico Interamericano del Cacao 5a, IICA. (5RTIC/DOC. 15-Fitop/2), , Turrialba, Costa Rica

ICCO (2017) ICCO quarterly bulletin of cocoa statistics, Vol. XLIII, No. 1, Cocoa year 2016/17. https://www.icco.org/statistics/production-and-grindings/production.html

Iwaro A, Butler D (2000) Germplasm enhancement for resistance to black pod and witches' broom diseases. In: Proceedings of the 13th international cocoa research conference: towards the effective and optimum promotion of cocoa through research and development, Kota Kinabalu, Sabah, Malaysia. Cocoa Producers' Alliance, pp 9–14

Iwaro AD, Sreenivasan TN, Umaharan P (1998) Cacao resistance to Phytophthora: effect of pathogen species, inoculation depths and pod maturity. Eur J Plant Pathol 104:11–15. https://doi.org/10.1023/A:1008622115731

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Johnson E, Phillips W, Bekele F, Zhang D, Schnell R (2007) Field guide to the UF clones of Costa Rica. In: Proceedings of the international cocoa producer's conference, San Jose, Costa Rica, Oct. 9–14 2007, pp 641–646

Johnson E, Rutherford M, Edgington S, Flood J, Crozier J, Cafá G, Buddie A, Offord L, Elliott S, Christie K (2017) First report of

Moniliophthora roreri causing frosty pod rot on Theobroma cacao in Jamaica. New Disease Reports 36:2–2

Keane P (2012) Horizontal or generalized resistance to pathogens in plants. In: Cumagun CJR (ed) Plant pathology. INTECH, Vienna, Austria., pp 317–352

Lanaud C, Kébé IS, Risterucci A-M, Clément D, N'Goran JAK, Grivet L, Tahi GM, Cilas C, Pieretti I, Eskes A, Despréaux D (1999) Mapping quantitative trait loci (QTL) for resistance to Phytophthora palmivorain T. cacao. In: Proceedings of the 12th international cocoa research conference, Salvador de Bahia, Brazil, 1996-11-17 / 1996-11-23, vol 12. Cocoa Producers' Alliance, none, pp 99–105

Lanaud C, Fouet O, Clément D, Boccara M, Risterucci AM, Surujdeo-Maharaj S, Legavre T, Argout X (2009) A meta–QTL analysis of disease resistance traits of Theobroma cacao L. Mol Breed 24:361–374. https://doi.org/10.1007/s11032-009-9297-4

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity 95:221–227. https://doi.org/10.1038/sj.hdy.6800717

Ling ASC, Kamil MJA, Chong KP, Ho CM (2017) Assessing the cocoa genotypes for resistance to black pod using the area under the disease-progress curve (AUDPC). Bulgarian Journal of Agricultural Science 23:972–979

Livingstone D, Royaert S, Stack C, Mockaitis K, May G, Farmer A, Saski C, Schnell R, Kuhn D, Motamayor JC (2015) Making a chocolate chip: development and evaluation of a 6K SNP array for Theobroma cacao. DNA Res 22:279–291. https://doi.org/10.1093/dnares/dsv009

Livingstone D, Stack C, Mustiga GM, Rodezno DC, Suarez C, Amores F, Feltus FA, Mockaitis K, Cornejo OE, Motamayor JC (2017) A larger chocolate chip—development of a 15K Theobroma cacao L. SNP array to create high-density linkage maps. Front Plant Sci 8. https://doi.org/10.3389/fpls.2017.02008

Mata-Quirós A, Arciniegas-Leal A, Phillips-Mora W, Meinhardt LW, Motilal L, Mischke S, Zhang D (2018) Assessing hidden parentage and genetic integrity of the "United Fruit Clones" of cacao (Theobroma cacao) from Costa Rica using SNP markers. Breed Sci 68:545–553. https://doi.org/10.1270/jsbbs.18057

McElroy MS, Navarro AJR, Mustiga G, Stack C, Gezan S, Pena G, Sarabia W, Saquicela D, Sotomayor I, Douglas GM, Migicovsky Z, Amores F, Tarqui O, Myles S, Motamayor JC (2018) Prediction of cacao (Theobroma cacao) resistance to Moniliophthora spp. diseases via genome-wide association analysis and genomic selection. Front Plant Sci 9:343. https://doi.org/10.3389/fpls.2018.00343

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185. https://doi.org/10.1093/nar/gkm321

Motamayor JC, Lachenaud P, e Mota JWdS, Loor R, Kuhn DN, Brown JS, Schnell RJ (2008) Geographic and genetic population differentiation of the Amazonian chocolate tree (Theobroma cacao L). PLoS ONE 3:e3311. https://doi.org/10.1371/journal.pone.0003311

Motamayor J, Mockaitis K, Schmutz J, Haiminen N, III D, Cornejo O, Findley S, Zheng P, Utro F, Royaert S, Saski C, Jenkins J, Podicheti R, Zhao M, Scheffler B, Stack J, Feltus F, Mustiga G, Amores F, Phillips W, Marelli J, May G, Shapiro H, Ma J, Bustamante C, Schnell R, Main D, Gilbert D, Parida L, Kuhn D (2013) The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. Genome Biol 14:r53

Motilal LA, Zhang D, Mischke S, Meinhardt LW, Boccara M, Fouet O, Lanaud C, Umaharan P (2016) Association mapping of seed and disease resistance traits in Theobroma cacao L. Planta 244:1265–1276. https://doi.org/10.1007/s00425-016-2582-7

Nyassé S, Cilas C, Herail C, Blaha G (1995) Leaf inoculation as an early screening test for cocoa (Theobroma cacao L.) resistance to

Phytophthora black pod disease. Crop Prot 14:657–663. https://doi.org/10.1016/0261-2194(95)00054-2

Olasupo FO, Adewale DB, Aikpokpodion PO, Muyiwa AA, Bhattacharjee R, Gutierrez OA, Motamayor JC, Schnell RJ, Ebai S, Zhang D (2018) Genetic identity and diversity of Nigerian cacao genebank collections verified by single nucleotide polymorphisms (SNPs): a guide to field genebank management and utilization. Tree Genet Genomes 14:32. https://doi.org/10.1007/s11295-018-1244-2

Osuna-Cruz CM, Paytuvi-Gallart A, Di Donato A, Sundesha V, Andolfo G, Aiese Cigliano R, Sanseverino W, Ercolano MR (2018) PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. Nucleic Acids Res 46:D1197–D1201. https://doi.org/10.1093/nar/gkx1119

Ouellette LA, Reid RW, Blanchard SG, Brouwer CR (2018) LinkageMapView-rendering high-resolution linkage and QTL maps. Bioinformatics 34:306–307. https://doi.org/10.1093/bioinformatics/btx576

Payne R, Welham S, Harding S (2015) A guide to REML in GenStat (18th Edition), 18th edn. VSN International, Hemel Hempstead, UK

Pertea M, Kim D, Pertea G, Leek JT, Salzberg SL (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie, and Ballgown. Nat Protoc 11:1650–1667. https://doi.org/10.1038/nprot.2016.095

Phillips-Mora W (1986) Evaluación de la resistencia de cultivares de cacao (Theobroma cacao L.) a Moniliophthora roreri (Cif. y Par.) Evans et al. Universidad de Costa Rica, San José .Centro Agronómico Tropical de Investigación y Enseñanza,

Phillips-Mora W (1996) Studies at CATIE on moniliasis resistance (Moniliophthora roreri (Cif. & Par.) Evans et al.). In: Proceedings of the international workshop on the contribution of disease resistance to cocoa variety improvement, Salvador, Bahía, Brazil. INGENIC, pp 111–117

Phillips-Mora W (2003) Origin, biogeography, genetic diversity and taxonomic affinities of the cacao (Theobroma cacao L.) fungus Moniliophthora roreri (Cif.) Evans et al. as determined using molecular, phytopathological and morpho-physiological evidence. University of Reading, Reading (RU),

Phillips-Mora W, Castillo J (1999) Artificial inoculations in cacao with the fungi Moniliophthora roreri (Cif. Par) Evans et al. and Phytophthora palmivora (Butl.) Butler. In: CATIE (ed) Actas. IV Semana CientíficaTurrialba. Logros de la investigacion para un nuevo milenio., Turrialba. Costa Rica. CATIE

Phillips-Mora W, Galindo JJ (1989) Métodos de inoculación y evaluación de la resistencia a Phytophthora palmivora en frutos de cacao (Theobroma cacao). Turrialba 39:488–496

Phillips-Mora W, Wilkinson MJ (2007) Frosty pod of cacao: a disease with a limited geographic range but unlimited potential for damage. Phytopathology 97:1644–1647. https://doi.org/10.1094/PHYTO-97-12-1644

Phillips-Mora W, Arciniegas-Leal A, Mata-Quirós A, Motamayor-Arias JC (2013) Catalogue of cacao clones selected by CATIE for commercial plantings. CATIE, Turrialba

Phillips-Mora W, Mata-Quiros A, Arciniegas-Leal A (2017) Generation of cacao clones with durable resistant against moniliasis/frosty pod rot (Moniliophthora roreri). In: International symposium on cacao research, Lima, Peru

Ploetz R (2016) The impact of diseases on cacao production: a global overview. In: Bailey BA, Meinhardt LW (eds) Cacao diseases: a history of old enemies and new encounters. Springer International Publishing, Cham, pp 33–59. https://doi.org/10.1007/978-3-319-24789-2_2

Pound FJ (1936) The completion of selection. In: Fifth annual report on cacao research 1935. Government Printing Office, Port of Spain, Trinidad, pp 7–16

Pound FJ (1943) Cacao and witches' broom disease (Marasmius perniciosus). Report on a recent visit to the Amazon territory of

Peru, September, 1942-February, 1943. Yuille, Port of Spain, Trinidad

Risterucci A, Paulin D, N'Goran J, Ducamp M, Lanaud C (2000) Mapping of quantitative trait loci (QTLs) for resistance to Phytophthora in Theobroma cacao L. INGENIC Newsletter:9–10

Romero Navarro JA, Phillips-Mora W, Arciniegas-Leal A, Mata-Quiros A, Haiminen N, Mustiga G, Livingstone Iii D, van Bakel H, Kuhn DN, Parida L, Kasarskis A, Motamayor JC (2017) Application of genome wide association and genomic prediction for improvement of cacao productivity and resistance to black and frosty pod diseases. Front Plant Sci 8:1905. https://doi.org/10.3389/fpls.2017.01905

Royaert S, Jansen J, da Silva DV, de Jesus Branco SM, Livingstone DS, Mustiga G, Marelli J-P, Araújo IS, Corrêa RX, Motamayor JC (2016) Identification of candidate genes involved in witches' broom disease resistance in a segregating mapping population of Theobroma cacao L. in Brazil. BMC Genomics 17:107. https://doi.org/10.1186/s12864-016-2415-x

Sánchez JA (1982) Reacción de cultivares de cacao a la inoculación artificial con Monilia roreri. Universidad de Costa Rica

Sánchez JA, Gonzalez LC (1989) Metodología para evaluar la susceptibilidad a moniliasis en cultivares de cacao (Theobroma cacao). Turrialba 39:461–468

Sánchez J, Brenes O, Phillips W, Enriquez G (1987) Methodology for inoculating pods with the fungus Moniliophthora (Monilia) roreri. In: Proc. 10th. cocoa res. conf., San Domingo, pp 467–471

SAS Institute (2016) SAS software version 9.4. SAS Institute, Cary, NC

Schnell R, Brown J, Kuhn D, Cervantes-Martinez C, Olano C, Motamayor J (2005) Why would we breed cacao in Florida? Proc Fla State Hort Soc 118:189–191

Schnell RJ, Brown JS, Kuhn DN, Cervantes-Martinez C, Borrone JW, Olano CT, Phillips W, Johnson E, Monteverde-Penso EJ, Motamayor JC, Lopes U, Amores F (2007) Current challenges of tropical tree crop improvement: Integrating genomics into an applied cacao breeding program. In: International symposium on biotechnology of temperate fruit crops and tropical species, 2007. International Society for Horticultural Science (ISHS), Leuven, Belgium, pp 129–144. https://doi.org/10.17660/ActaHortic.2007.738.11

Sotomayor F (1965) Estudios preliminares sobre la resistencia de algunos clones de cacao a la Moniliasis provocada por la inoculación artificial. Universidad de Guayaquil

Spence JA (1961) Probable mechanism of resistance of varieties of cocoa to black pod disease caused by Phytophthora palmivora (Butl.) Butl. Nature 192:278–278

Suárez-Capello C (1999) Monilia pod rot resistance in Ecuador. In: International workshop on the contribution of disease resistance to cocoa variety improvement, Salvador, BA, Brazil. INGENIC, pp 119–121

Surujdeo-Maharaj S, Sreenivasan TN, Motilal LA, Umaharan P (2016) Black pod and other Phytophthora induced diseases of cacao: history, biology, and control. In: Bailey BA, Meinhardt LW (eds) Cacao diseases: a history of old enemies and new encounters. Springer International Publishing, Cham, pp 213–266. https://doi.org/10.1007/978-3-319-24789-2_7

Tan GY, Tan WK (1990) Additive inheritance of resistance to pod rot caused by Phytophthora palmivora in cocoa. Theor Appl Genet 80: 258–264. https://doi.org/10.1007/BF00224396

Utro F, Haiminen N, Livingstone D, Cornejo O, Royaert S, Schnell R, Motamayor J, Kuhn D, Parida L (2013) iXora: exact haplotype inferencing and trait association. BMC Genet 14

Van Ooijen J (2018) JoinMap ® 5, software for the calculation of genetic linkage maps in experimental populations of diploid species. Kyazma B.V, Wageningen, Netherlands

Van Ooijen JW, Jansen J (2013) Genetic mapping in experimental populations. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139003889

VSN International (2017) Genstat for Windows 19th Edition. VSN International, Hemel Hempstead, UK

Whitlock BA, Baum DA (1999) Phylogenetic relationships of Theobroma and Herrania (Sterculiaceae) based on sequences of the nuclear gene vicilin. Syst Bot:128–138

Zarrillo S, Gaikwad N, Lanaud C, Powis T, Viot C, Lesur I, Fouet O, Argout X, Guichoux E, Salin F, Solorzano RL, Bouchez O, Vignes H, Severts P, Hurtado J, Yepez A, Grivetti L, Blake M, Valdez F (2018) The use and domestication of Theobroma cacao during the mid-Holocene in the upper Amazon. Nat Ecol Evol 2:1879–1888. https://doi.org/10.1038/s41559-018-0697-x

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468