Tree Genetics & Genomes (2020) 16: 89
https://doi.org/10.1007/s11295-020-01479-3

ORIGINAL ARTICLE

A reference high-density genetic map of *Theobroma grandiflorum* (Willd. ex Spreng) and QTL detection for resistance to witches' broom disease (*Moniliophthora perniciosa*)

Pierre Mournet[1,2] · Paulo Sérgio Beviláqua de Albuquerque[3] · Rafael Moysés Alves[4] · Joseilde Oliveira Silva-Werneck[5] · Ronan Rivallan[1,2] · Lucilia Helena Marcellino[5] · Didier Clément[1,2]

Received: 20 August 2020 / Revised: 2 November 2020 / Accepted: 5 November 2020 / Published online: 22 November 2020

Abstract

Theobroma grandiflorum (Willd. ex Spreng.) Schum. is a tree originating in the Amazon forest. In Brazil, it is known as cupuassu and, like *Theobroma cacao*, belongs to the family of the Malvaceae. The pulp is used by the food industry and the seeds, which are rich in fatty acids, are used in cosmetics. As for cacao, cupuassu cultivation is affected by the witches' broom disease (WBD), caused by *Moniliophthora perniciosa* (*M. perniciosa*). Despite its economical value, cupuassu culture lacks genetic information, compromising its plant breeding program development. The aim of this research was to increase molecular knowledge about WBD resistance. The progeny of 168 individuals obtained by crossing two contrasting cupuassu clones (174 and 1074, resistant and susceptible to WBD, respectively) were evaluated. The first consensus genetic map consisting of 1438 markers was produced using genotyping by sequencing (GBS) and genome sequencing of the neighboring species *T. cacao* (Criollo B97) as reference. The analysis of synteny between *T. cacao* and *T. grandiflorum* showed that the average homology between the linkage groups of the two species was 97.2%. A resistance quantitative trait locus (QTL) was identified on chromosome 6 of resistant parent 174 at marker 6M1252980. The phenotypic data associated with that QTL corresponded to observations after natural and artificial infection with *M. perniciosa*. The biological process associated with this QTL may play a role in resistance and susceptibility to *M. perniciosa*.

Keywords Cupuassu · GBS · QTL · Witches' broom disease · *Moniliophthora perniciosa*

Introduction

Theobroma grandiflorum (Willd. ex Spreng.) Schum. is a fruit tree originating from the Amazon forest, belonging to the Malvaceae family. It is a diploid with 2n = 20, like *Theobroma cacao* (Kuhn et al. 2010; da Silva et al. 2017). Like other fruit trees of interest, *T. grandiflorum*, known as cupuassu in Brazil, has been dispersed and domesticated by the native populations of Amazonia (Clement 1999). Since the 1970s, it has been grown on a larger scale on small farms mostly located in the states of Pará, Amazonas, and Rondônia (Alves et al. 2010, 2013). The commercial value of cupuassu comes from the pulp surrounding the seeds, and from the fat extracted from them. The pulp is mostly used by the food industry to make juices, sweets, and ice creams (Costa et al. 2003; De Oliveira and Genovese 2013; Pugliese et al. 2013). The seeds yield a fine fat containing various compounds, amino acids, and vitamins, which are of great interest for the cosmetics industry (Gilabert-Escrivá et al. 2002; Costa et al. 2016). *Embrapa Amazônia Oriental*, located in Belém in the state of Pará (Brazil), is in charge of the cupuassu breeding program. After accession collection, formation of ex situ collections, and characterization, a genetic

Communicated by D. Chagné

✉ Didier Clément
didier.clement@cirad.fr

[1] CIRAD, UMR AGAP, F-34398 Montpellier, France

[2] AGAP, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France

[3] Comissão Executiva do Plano da Lavoura Cacaueira, BR 316, km 17, CP 46, Marituba, Pará CEP 67105-970, Brazil

[4] Empresa Brasileira de Pesquisa Agropecuária, Embrapa Amazônia Oriental, Belém, PA 66095-903, Brazil

[5] Empresa Brasileira de Pesquisa Agropecuária, Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF 70770-917, Brazil

improvement program was launched, and a selection scheme was introduced, enabling the dissemination of hybrid seeds and clones to farmers (Maia et al. 2011; Alves et al. 2017). Cupuassu cultivation is affected by the witches' broom disease—WBD (Alves and Resende 2008) caused by *Moniliophthora perniciosa*, a hemibiotrophic basidiomycete (Aime and Phillips-Mora 2005). Meristems located on the branches or floral cushions are infected by the pathogen, which causes swelling of the rod known as brooms. The pathogen also attacks the cells of the fruit cortex, subsequently making the seeds unusable. WBD is the main limiting factor for cupuassu production, as it is also for cacao cultivation in Brazil, where the main objective of selection is to obtain plants resistant to this disease (Lopes et al. 2011).

In relation to cacao WBD, studies have been undertaken on the diversity of resistance sources (Albuquerque et al. 2009; Pires et al. 2012; Gramacho et al. 2016; Neto et al. 2018), genetic mapping, and detection of resistant quantitative trait locus (QTL) to WBD (Brown et al. 2005; Santos et al. 2014; Royaert et al. 2016). Some functional genomics studies based on understanding of the cocoa transcriptome of the interaction between *T. cacao* and *M. perniciosa* have also been carried out (Gesteira et al. 2007; Argout et al. 2008; Leal et al. 2010; Teixeira et al. 2014). Cacao genome sequencing studies (Argout et al. 2011; Motamayor et al. 2013; Argout et al. 2017) have made a major contribution to genetic and genomic studies, and have been extensively supporting cacao breeding program. A new version v2 of the Criollo B97 genome published by Argout et al. (2017) provides virtually total coverage of the cacao genome. In addition, and concerning to the pathogen, a *M. perniciosa*-sequencing study has also been undertaken (Montego et al. 2008; Barbosa et al. 2018).

Genetic analyzes on cupuassu using molecular markers began with microsatellites primarily derived from cacao genomic libraries and transferred to cupuassu (Alves et al. 2006; Alves et al. 2007; Alves et al. 2017). Subsequently, microsatellites from cupuassu genomic libraries were used (unpublished data), followed by microsatellites derived from expressed sequences (Ferraz dos Santos et al. 2016a, 2016b; Nascimento et al. 2019). A cytogenetic approach was also carried out for different species of *Theobroma*, including *T. grandiflorum* (Da Silva et al. 2017). More recently, differential expression data were provided by RNAseq analyses of cupuassu apical shoots of resistant and susceptible clones (174 and 1074, respectively), inoculated with *M. perniciosa* (Falcão et al. 2018). The differential gene expression previously detected by RNAseq was analyzed by RT-qPCR from cupuassu apical meristems of the genotypes 174 and 1074 inoculated with *M. perniciosa* (Silva et al. 2020). This analysis showed a higher level of expression in 3 genes encoding for PR proteins with chitinase and osmotin biological activity: TgPR3, TgPR5, and TgPR8, and revealed the involvement in the process of resistance to the pathogen (Silva et al. 2020). These data suggest that this gene could be involved in recognizing mechanism of the pathogen. These two clones were also the parents of the cross used in our study, which sought to generate the first high-density genetic mapping of *T. grandiflorum* using GBS. The other results presented in this study cover an analysis of synteny between *T. cacao* and *T. grandiflorum*, the detection of a QTL of resistance to WBD, and an analysis of the genome region linked to the latter.

Material and methods

Population and plant material

Clones 174 and 1074 were the female and male parent, respectively, of the progeny (174 × 1074). The two parents came from the *Embrapa Amazônia Oriental* collection, located in Belém (Pará, Brazil). Parent 174, called *Coari*, is known to be resistant to WBD, while clone 1074 is susceptible. The progeny of 168 individuals was planted with a first part (102 individuals) in April 2011 and 66 in 2012 at the experimental station of CEPLAC (*Comissão Executiva do Plano Lavoura Cacaueira*), located in Marituba (Pará, Brazil). In addition, the two parental clones, 174 and 1074, were planted in a plot located alongside the progeny plot, where they were reproduced by cloning to provide 45 replicates.

DNA extraction, GBS libraries, and sequencing process

Leaves were collected from the 168 individuals of the 174 × 1074 progeny, and DNA was extracted according to a protocol using mixed alkyl trimethylammonium bromide (MATAB) buffer already described for the isolation of genomic DNA (Risterucci et al. 2000). DNA samples were quantified with a Fluoroskan Ascent FL fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Genomic DNA quality was checked by agarose gel electrophoresis. Genomic DNA concentration was adjusted to 20 ng/μL, and PstI/MseI GBS libraries were prepared following the protocol described by Cormier et al. (2019). Single-end sequencing of 150 base-pair reads was performed in a single lane on an Illumina HiSeq 3000 system at the GeT-PlaGe platform in Toulouse, France.

SSR genotyping

Eight SSR (Tg11, Tg13, Tg20, Tg39, Tg54, Tg62, Tg65, Tg72) from the cupuassu genomic library (unpublished data) and 2 EST-SSR (c180 and c4949) from EST-sequences (Ferraz dos Santos et al. 2016a) were used. The sequences of the primers are given in Table 1 (additional file). PCR reactions were performed as simplex experiments in a volume of 10 μL containing 1 X reaction buffer (100 mM Tris-HCl,

pH 9.0; 100 mM KCl; 80 mM $(NH4)_2SO_4$; 1% Triton X-100), 200 μM dNTP (Jena Bioscience GmbH, Jena, Germany), 0.5 mM $MgCl_2$, 0.08 μM forward primer with a M13 tail at the 5′-end, 0.1 μM reverse primer, 0.1 μM of fluorescent-labeled M13-tail (6-FAM, NED, VIC, or PET from Applied Biosystems, Foster City, CA, USA), and 0.6 U Taq DNA Polymerase (Taq'Ozyme OZYA001 from Ozyme, Montigny-le-Bretonneux, France), 5 ng of template DNA. A touchdown cycling program was used as follows: an initial denaturation at 94 °C for 5 min, followed by 10 cycles at 94 °C for 30 s, 55 °C for 60 s (0.5 °C decrease at each cycle), 72 °C for 1 min, followed by 25 cycles at 94 °C for 45 s, 50 °C for 1 min, 72 °C for 1 min, and a final extension at 72 °C for 30 min. Fluorescently labeled PCR products were then organized in five pools for electrophoresis, using respectively 2 μL of products labeled with 6 FAM, 2 μL of those with VIC, 2.5 μL of those with NED, and 3.5 μL of those with PET, and completed to 20 μL with high purity water. Two microliters of this solution was taken and added to 10 μL of Hi-Di formamide and 0.12 μL of GeneScan 600 LIZ size standard (Applied Biosystems). Migration of PCR products was performed on an ABI 3500xL Genetic Analyzer (Life Technologies, Carlsbad, CA, USA). Alleles were scored using the GeneMapper v.4.1 software (Applied Biosystems).

SNP calling and filtering

The GBS analysis pipeline implemented in the Tassel GBS version 5.2.29 (Glaubitz et al. 2014) was used to call SNP according to the parameters listed in the Script 1 (additional file). SNP and progenies were filtered using the following filters: (1) indel variant sites were removed; (2) genotype calls based on < tenfold coverage, < threefold coverage for the minor allele, and maximum depth 500 were converted to missing data; (3) variant sites with > 20% missing genotype calls were filtered out; (4) minor allele frequencies per site > 10%; and (5) only biallelic sites were kept. The progenies were selected based on their Mendelian segregation ratio with 1:1 if segregating only in one parent and 1:2:1 if segregating in both parents. Markers with a significant segregation deviation based on a χ^2 test with a maximum P value of 1e-16 were eliminated using the VcfPreFilter.1.0.py (python2) program with the default parameters. Both programs are part of the Vcf-Hunter package (Garsmeur et al. 2018) available at https://github.com/South Green Platform/VcfHunter.

Consensus and parental genetic map construction

Linkage analysis and map constructions were performed separately for each parent according to the cross-pollinated (CP) model using the JoinMap version 5 (Plant Research International BV, Wageningen, Netherlands). The "lm × ll" and "nn × np" SNP segregation patterns were used for the female (174) and male parental (1074) maps, respectively. The segregation pattern from 8 SSR (cupuassu genomic library) and 2 EST-SSR was integrated into the genotyping matrix. The marker loci were attributed to linkage groups using the JoinMap grouping module with a logarithm of odds (LOD) score of 7.0. Marker order and genetic distance (in cm) were calculated using a regression mapping algorithm with the following parameters: Kosambi's mapping function, recombination frequency 0.45, LOD threshold of 1.0, and regression algorithm with 3 ordering rounds. Linkage group (LG) numbering and orientation were attributed to best fit the *T. cacao* chromosomes. Maps were plotted using SpiderMap software (JF Rami,

Table 1 Distribution of 1438 SNP over 10 linkage groups of *Theobroma grandiflorum* consensus map

Linkage group	Number of SNP	Genetic length (cm)	Marker interval (cm) *	Max gap (cm)	Number of gaps > 5 cm	Physical length cacao genome (Mb) **
LG 1	179	124.2	0.859	14.022	1	37.3
LG 2	173	171.282	0.996	13.249	2	41.2
LG 3	147	164.612	1.127	10.735	5	36.4
LG 4	148	132.777	0.903	5.828	3	31.9
LG 5	187	152.036	0.817	6.951	4	39.4
LG 6	130	143.649	1.114	12.06	2	26.3
LG 7	119	108.061	0.916	5.587	1	21.6
LG 8	75	101.623	1.373	3.928	0	19.6
LG 9	147	169.03	1.158	6.828	2	38.6
LG 10	133	113.246	0.858	6.209	1	21.8
	1438	1380.516	0.987	14.022	21	314.1

* Mean distance between SNPs

** Length of *T. cacao* chromosome (Argout et al. 2017)

unpublished). The *T. grandiflorum* consensus map was constructed using the same parameters as the parental maps, including the "hk × hk" segregation patterns that were used for both parents.

The consensus map was compared to the *T. cacao* reference genome (Argout et al. 2017). The *T. cacao* genome was divided into fragments cutting halfway between SNP included in our *T. grandiflorum* reference genetic map. The resulting genomic fragments were then reordered according to the SNP positions in the *T. grandiflorum* reference map developed in this study. Synteny was visualized using a Circos approach via the circlize R package 0.4.3 (Gu 2014).

Witches' broom disease observations

The phenotypic data on witches' broom disease were obtained from 11 evaluations carried out under natural infection conditions over a period stretching from 2014 to 2019. Brooms from infected meristems on branches belonging to the crown of the tree were called "Terminal Brooms" (TB). Some artificial meristem inoculations were carried out in the field using a strain of *M. perniciosa* from dry brooms, sampled at the CEPLAC station in Belém, and placed under artificial conditions to obtain basidiospores (Frias et al. 1995). For each individual in the progeny planted in the field, four meristems were inoculated with 30 μL of a spore suspension at a concentration of 1×10^5 spores/mL. The inoculated parts were protected for 24 h under humid conditions, and the presence or absence of brooms was observed 60 days after infection. Both types of observations are illustrated in Fig. 1 (additional file). The trunk circumference was measured around 1 m from the ground to obtain an estimation of the vigor of the progeny individuals at the beginning and end of disease observations.

Statistical analyses and QTL detection

Descriptive statistical analyses and phenotypic data modeling were done with XLSTAT software (Addinsoft). QTL detection analyses were carried out with MapQTL 6 (Van Ooijen 2006). QTL significance, LOD, was obtained from the permutation test proposed by MapQTL and based on the empirical method established by Churchill and Doerge (1994) to determine the significance of the LOD threshold. The significance LOD thresholds, with an alpha risk of 0.05, were calculated from 10,000 permutations to check the stability of the threshold. The method used for QTL detection was Simple Interval Mapping (SIM) by the multiple-QTL model, MQM, making it possible to refine the analysis and the positioning of the significant QTL detected in IM and using cofactors, which are the markers located at the peak level of the LOD or close to it.

Results and discussion

GBS analysis and SNP filtering on the F1 population

A progeny of 168 individuals and the 174 × 1074 contrasting parents were genotyped by the GBS approach. We obtained 954,054,141 Illumina raw reads available in the NCBI SRA (Sequence Read Archive), under the study accession number PRJNA609621. In all, 1,736,489 tags were detected with the Tassel GBS pipeline and aligned to 901,020 (51.89%) for a unique position, 344,172 (19.82%) for multiple positions, and 491,297 (28.29%) could not be aligned to the cacao genome v2. Tag mapping and polymorphism calling identified 742,999 polymorphic sites. From the 5370 high-quality SNP filtered, 1937 (36.1%) SNP were classified as female parent segregation ("174"; <lm x ll>), 1773 (33.0%) as male parent segregation ("1074"; <nn x np>), 1087 (20.0%) SNPs as heterozygous segregation in both parents (<hk x hk>), and 573 (10.7%) SNP as unknown. The final dataset for genetic mapping included 141 individuals of the progeny; 27 individuals were removed from the matrix due to the presence of a lot of missing data.

Parental genetic maps

Parental genetic maps of 174 and 1074 were established only and respectively from the marker codes, ll lm, and nn np and from 141 individuals. After the elimination of SNP during the different phases of construction, we obtained the parental maps of both 174 and 1074 clones, which are presented in the Table 2 (additional file). Because of the strong linkages revealed by the pairwise recombination frequencies and LOD scores (Fig. 2) (additional file), linkage groups were confidently defined for each parent. The number of linkage groups per parental map was 11 for 174 and 10 for 1074. It was found that LG 4 of parent 174 was split into two parts (4a and 4b). The parental genetic maps of both 174 and 1074 were equivalent with a total length of 1207 and 1222 cm, respectively (Table 2 additional file).

Consensus genetic map

The consensus genetic map constructed from 141 individuals was obtained from 606 markers (ll, lm) for parent 174, 673 markers (nn, np) for parent 1074, and 159 markers (hk x hk) as bridge markers. Ten SSR were also mapped. The results obtained for each linkage group and the drawing of the map are presented in Table 1 and Fig. 1, respectively. The consensus genetic map had a total length of 1380 cm and the linkage analysis was carried out using 1438 SNP. The linkage analysis showed 10 linkages groups, which corresponded to the reported haploid chromosome number in *T. grandiflorum*. Map integration was accurate, as revealed by the good collinearity between homolog linkage groups from the different parental maps (Fig. 2).

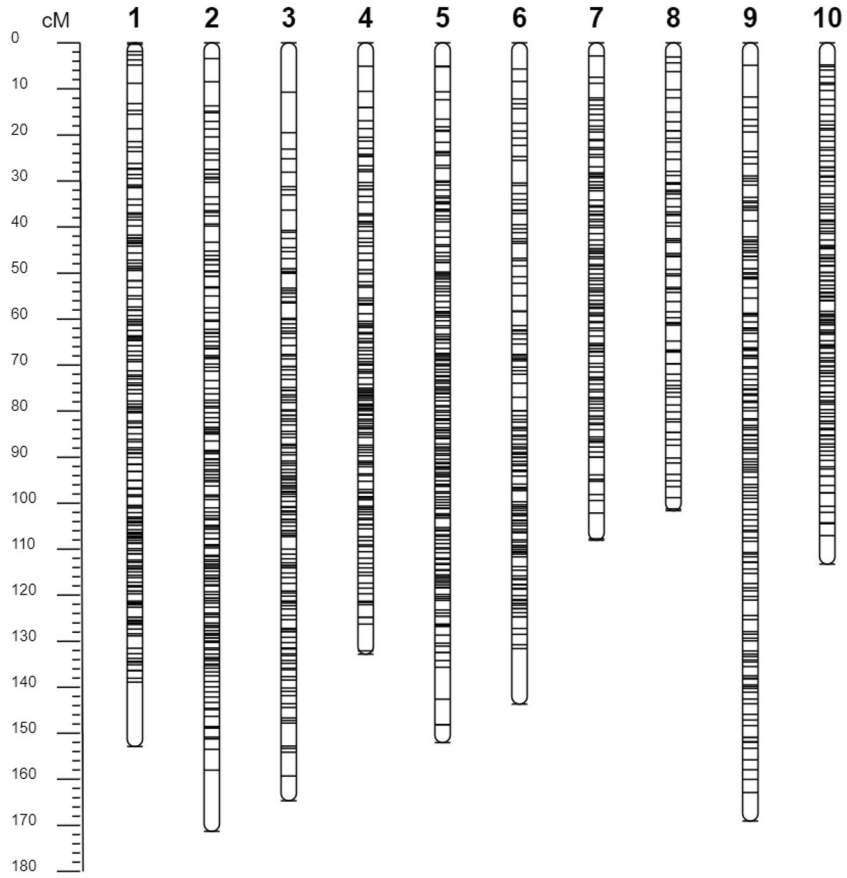

Fig. 1 *Theobroma grandiflorum* consensus genetic map from 141 individuals and 1438 SNP over 10 linkage groups. Bars on each linkage group represent SNP markers. *Y*-axis, genetic distance in cm (Kosambi mapping function)

The length of each group ranged from 101.623 (LG8) to 171.282 cm (LG2). Linkage group 5 showed the highest average density and the highest number of markers mapped (187), whereas linkage group 8 showed the lowest density (75). The SNP identified by GBS were uniformly distributed across the genetic map with an average distance between markers of 0.987 cm/SNP. We observed 21 regions with a low marker density and gaps greater than 5 cm. The marker saturation of the map is shown in the Figure 3 (additional file).

Synteny between *T. grandiflorum* and *T. cacao*

Most of the SNP identified using the cacao genome were localized in the correct linkage group after linkage analysis.

Table 2 Divergence between *T. grandiflorum* (cupuassu) linkage groups and the *T. cacao* chromosomes

Linkage group	Number of SNPs	Number of SNP present in a different cacao chromosome	% of mismatch	Comment on the position on the cacao chr *
LG 1	179	5	6.3	chr 2, 3, & 8
LG 2	173	2	2.7	chr 3 & 5
LG 3	147	3	2.0	3 markers in cluster of chr 5
LG 4	148	0	0	
LG 5	187	2	1.0	chr 2 & 9
LG 6	130	1	0.8	chr4
LG 7	119	5	4.2	3 markers in cluster of chr 6, another of chr 6 & chr 8
LG 8	75	2	2.7	3 markers in cluster of chr 7
LG 9	147	2	1.4	2 markers in cluster of chr 8
LG 10	133	12	9.0	2 markers in cluster of chr 4, 2 markers in cluster of chr 5, 2 markers in cluster of chr 6 & 2 markers in cluster of chr 7
	1438	34	2.4	

* Description on the arrangement of the conflicting markers between each cupuassu linkage group and the cocoa physical map

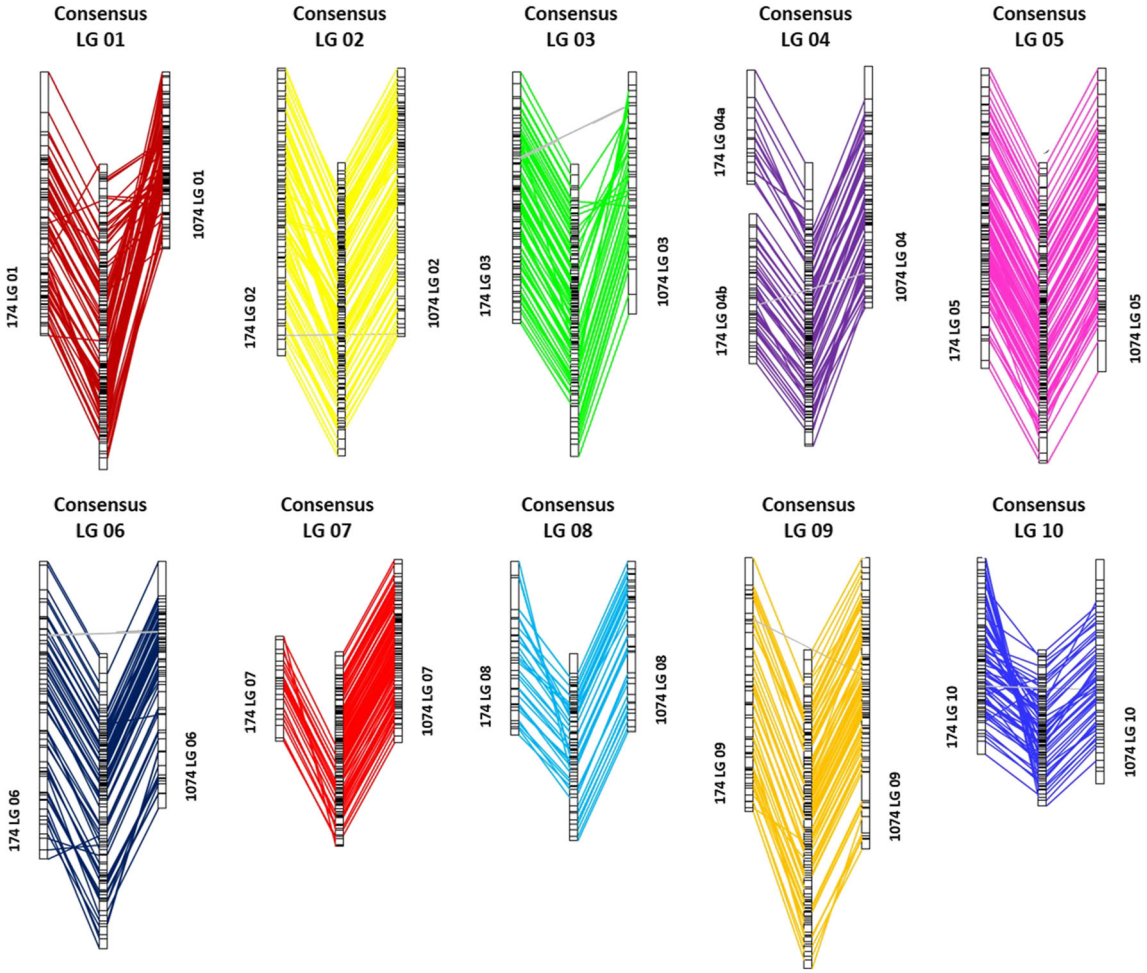

Fig. 2 *Theobroma grandiflorum* map integration with the homolog linkage groups from the different parental map. The scale and orientation of the linkage groups are the same as used to represent for the consensus map (Fig. 1)

However, some SNP did not follow the same physical order relative to the chromosomes of cacao v2. The mean homology between the *T. grandiflorum* and *T. cacao* linkage groups was 97.2% when computed as the percentage of SNP located in a *T. grandiflorum* linkage group and in its *T. cacao* homolog. All linkage groups of *T. grandiflorum* showed hits mainly with one chromosome of *T. cacao* and shared the same base chromosomes number of 10.

In comparison, the reference consensus genetic map of *T. cacao* was established with 1262 codominant markers (mainly from SNP) with a total length of 733 cm (Allègre et al. 2012). From the ratio between the length of each linkage group and the total length of each map, it can be seen in the Figure 4 (additional file) that the linkage groups of the two species of *Theobroma* were similar.

In an additional analysis of transposon fingerprinting using low coverage, whole-genome shotgun sequencing in cacao (*T. cacao*) and related species (Sveinsson et al. 2013), the author concluded that *T. grandiflorum* has a similar transposable elements (TE) composition to *T. cacao*. Lastly, it has been demonstrated by cytogenetic and flow cytometry approaches that there is stability in genome size with a karyotype 2n = 20 for species of the genus *Theobroma* (Da Silva et al. 2017).

However, some SNP did not follow the same physical order or relative position as the chromosomes of cacao v2. Based on the SNP positions in our *T. grandiflorum* consensus genetic map compared to SNP positions in the *T. cacao* genome, the *T. cacao* genome was reordered: (i) to highlight possible chromosome rearrangements between the two species and (ii) to estimate the relationship between physical and linkage map distances in *T. grandiflorum* (Fig. 3).

A detailed analysis of macrosynteny between homologous chromosomes showed evidence of internal rearrangements. LG 10 displayed the biggest rearrangement, presenting 12 markers originating from other chromosomal regions in the cocoa tree genome. Conversely, LG 4 seemed to be the most conserved. It should also be noted that some of these markers present on another linkage group were most often inserted in a cluster and were a few kb apart, as illustrated in Table 2. This suggests that some translocation and inversion events occurred between these two genomes. However, the fine

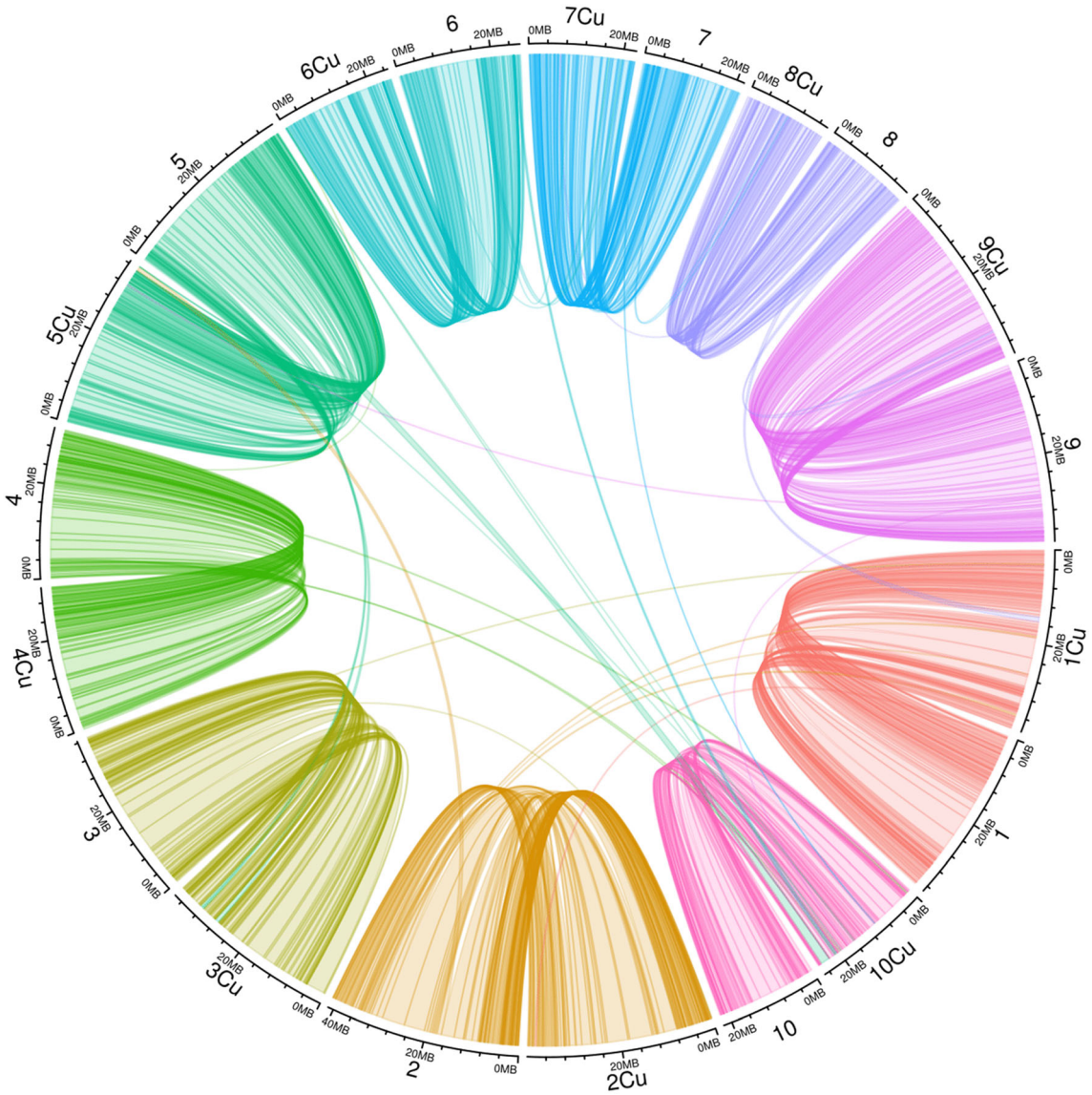

Fig. 3 Synteny between *T. grandiflorum* and *T. cacao*. The suffix "Cu" stands for the *T. grandiflorum* (cupuassu) linkage group

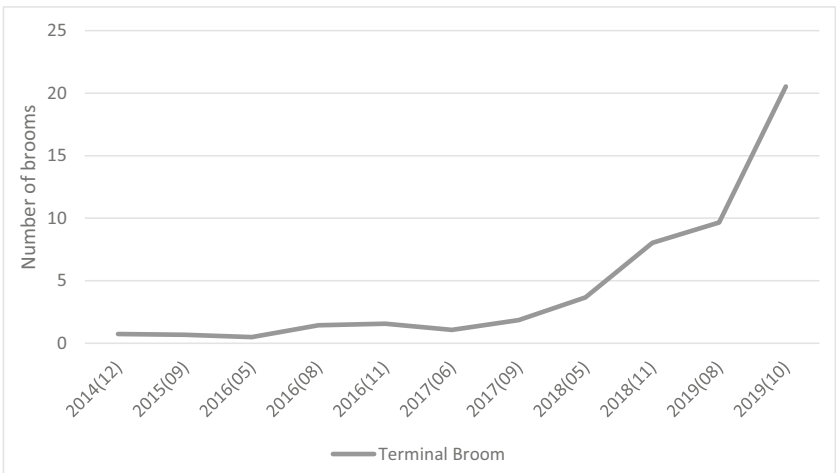

Fig. 4 Witches' broom disease evolution observed in127 individuals of the progeny 174 × 1074 from the number of terminal brooms determined over 11 evaluations carried out between 2014 and 2019. The number in parenthesis corresponds to the month of the evaluation year

mapping of such regions will require a larger set of molecular markers, or a more specific molecular analysis (mate pair library sequencing or long read sequencing). For this reason, these local inversions cannot be described with certainty, since artifactual origins were also quite probable. Several studies also concluded that collinearity at the intergeneric level was highly conserved between genetic maps and physical maps on related species, with the presence of reshuffles in the genome regions (Ollitrault et al. 2012; Cormier et al. 2019; Salgon et al. 2017; Chaintreuil et al. 2016). Comparative mapping between genomes can reveal the genetic basis of conservation and differences between related species. Anchoring of the linkage map of field cross to the *T. cacao* genome revealed more regions of conserved synteny, which will play a substantial role in the discovery of novel genes or alleles that can be utilized in the domestication and molecular breeding of *T. grandiflorum*. To our knowledge, our results are the first direct syntenic analysis of both species.

Statistical analyses from phenotypic observations related to WBD

Terminal brooms (TB) were counted in the plot containing 45 replicates of each of the two parents, 174 and 1074, in November 2018, and the result is illustrated in Figure 5 (additional file), which shows a very clear difference between the resistant parent 174 (average TB = 0.11) and the susceptible parent 1074 (average TB = 36.22).

The average broom numbers per individual were established from 11 sets of observations carried out between 2014 and 2019. The histogram of these data is presented in Figure 6 (additional file) and the progression of the TB evaluation over that period is presented in Fig. 4. It can be seen that the number of type TB brooms increased substantially from September 2017, when the disease began. The vigor of the individuals was estimated from the trunk circumference measured a year before observations began, in 2013, then at the end of observations in 2019. The two vigor distributions for these two periods are shown in Figure 7 (additional file). It can be seen that there was less scatter from the average vigor value in 2019 compared to 2013, when there was substantial heterogeneity for this criterion due to the offset in the field installation (2011 to 2012). Pearson's correlation established from the vigor of the trees in 2019 and the mean of broom numbers obtained for the 2014 to 2019 period (TB14–19) was 0.082 with $P = 0.363$. If the correlation was calculated from the average TB number between 2017 and 2019 (TB17–19), i.e., over the period corresponding to the spread of the disease, Pearson's correlation was then 0.004 with $P = 0.975$. Based on these results, it can be considered that there was no vigor effect on the number of brooms.

Artificial inoculations were carried out on 99 individuals of the progeny and the broom percentages obtained 60 days after inoculation showed that 2/3 of the individuals did not display any brooms, as shown in Figure 8 (additional file). Pearson's correlation, established from observations of the number of brooms (TB14–19) arising from natural infections and the percentage of brooms (TB) obtained after artificial infection, was 0.456 with $P < 0.0001$. This result showed that there was a similar response of the individuals in terms of resistance, independently of the type of infection. This suggests (i) a similarity of strains for the two infection methods and (ii) the same biological process in play to counter the pathogen.

QTL detection of WBD resistance

The QTL detection analyses based on observations under natural infection conditions (TB14–19) were carried out on 127 individuals out of the 141 used to produce the genetic maps, with simple interval mapping (SIM) and composite interval mapping, using the MQM procedure. A QTL was identified on chromosome 6 of parent 174; the LOD peak was positioned on marker 6M1252980, for a LOD value of 3.75. The significance threshold value, established by the mutation test, was 2.5 for 0.95, and the percentage of phenotypic variance explained by the QTL was 12.8. A second QTL detection analysis was carried out with the 99 individuals used for the artificial inoculations and for the broom percentage variable (%TB). The same QTL was identified; the LOD values at the QTL peak for %TB and the percentage of phenotypic variance explained were 6.12 (with a significance threshold of 3.2 for

Table 3 QTL of resistance to *M. perniciosa* detected on linkage group 6 in the 174 genetic map and involving two SNP: 6M1252980 and 4M27334640

Trait*	Group	RP cm	Peak cm	LP cm	Locus marker	LOD value	Mapping method	Variance explained	Additive effect
TB14–19	6	17.55	21.06	28.48	6M1252980	3.75	MQM	12.8	1.544
TB14–19	6	17.55	19.22	28.48	4M27334640	3.41	MQM	16.48	1.487
%TB	6	17.55	21.06	28.48	6M1252980	6.12	MQM	28.4	15.709
%TB	6	17.55	19.22	28.48	4M27334640	5.37	MQM	22.5	14.139

* TB14–19 is the average of terminal brooms, observed between 2014 and 2019, due to natural inoculation; %TB refers to terminal brooms obtained after artificial inoculation. RP is the right position of the QTL interval, peak is a position of the QTL with the maximum LOD value, and LP is left position of the QTL interval

0.95) and 28.4, respectively. These results are shown in Table 3. Marker 4M27334640, corresponding to a marker located in the distal section of cacao tree chromosome 4, was found to be mapped in the QTL region of chromosome 6 at 1.7 cm from marker 6M1252980. The chromosomal region associated with marker 4M27334640 was part of the QTL confidence interval. For this marker, the LOD values at the QTL peak and the percentage of phenotypic variance explained by the QTL were 3.41 and 5.35, respectively (Table 3).

This suggested that it was probably the same strain of *M. perniciosa* causing the symptoms obtained from the two types of infections (TB14–19 and %TB). In addition, when considering the genome region concerned by the QTL, the reference and substitution alleles found for SNP 6M1252980 and 4M27334640 were C/A and T/C, respectively. After classing the individuals according to the type of allele and in relation to the two variables analyzed, the AC individuals for 6M1252980 and TC for 4M27334640 could be considered resistant and the homozygous CC individuals as susceptible. Based on the confidence interval of the MQM method, the QTL of resistance to *M. perniciosa* revealed on parent 174 occurred in the distal position of chromosome 6, between markers 6M488189 and 6M4175947 on 9 marker and a distance of 13.92 cm. The configuration of the haplotypes for the 9 markers of this region of the genome, linked with the behavior of the resistant/susceptible individuals, is presented in Table 4.

Analysis of the WBD QTL region on the cacao physical map

The strategy of detecting SNP polymorphisms using the reference cacao genome enabled us to directly link the detected QTL zones in relation to the cacao physical map. Several sources of resistance to WBD have been identified in cacao trees (Pires et al. 2012; Gramacho et al. 2016; Neto et al. 2018); among them, the Scavina-6 resistance source has been one of the most widely used as a parent in various cacao hybrid progenies distributed to farmers. Various QTL of resistance to WBD have been found (Brown et al. 2005; Santos et al. 2014; Royaert et al. 2016). These studies showed that the resistance to WBD is often linked to few QTL at the outset of infection (Brown et al. 2005) and that more QTL can be detected after several years of plant-pathogen interaction (Santos et al. 2014; unpublished data). However, among the significant WBD-resistant QTL published in cocoa, none was observed on chromosome 6. In addition, based on observations carried out by CEPLAC (Marituba, Pará, Brazil), no evidence of cross-infection has yet been found between *M. perniciosa* isolates from cupuassu and cocoa (CEPLAC, Departamento Especial Da Amazônia, report 1986). It seems that for these two species of *Theobroma*, there are specificities in the host-

Table 4 Resistant and susceptible haplotypes description over nine markers covering 34.8 cm of the QTL for resistance to *M. perniciosa* from the 174 × 1074 progeny

Markers*	6M1495733	6M154484	6M488189	4M27334640	6M1252980	6M4175947	6M2460094	6M2509782	6M3578886	Effect if by haplotype	WBD data**
cm	0	1.84	14.55	19.22	21.03	28.48	29.44	29.83	34.81		
Allele	C/T	A/G	A/C	C/T	A/C	A/G	C/T	C/T	C/T		
R				CT	AC					8	0.36
	CT	AG	AC	CT	AC	AG	CT	CT	CC	6	
S				CC	CC					12	8.14
	CC	AA	CC	CC	CC	AA	CC	CC	CC	3	
	CC	AA	CC	CC	CC	AA	CC	CC	CT	9	

R resistant and *S* Susceptible haplotypes

* Nine markers without missing data over 38 individuals

** WBD data: phenotypic data on witches' broom disease obtained from an average of terminal brooms over 11 evaluations for a period stretching from 2014 to 2019 and a gray area corresponding to variable allelic composition

pathogen interactions; the defense and recognition mechanisms used against *M. perniciosa* are different.

An investigation of the zone of the WBD QTL detected from the cupuassu genetic map revealed the existence of 398 genes (search for chr6:450000..4180000 on Cocoao Genome Hub: https://cocoa-genome-hub.southgreen.fr/) in a region of around 3.7 Mb. In this interval, 21 putative disease resistance genes are listed on the cocoa genome without giving any indication of a real role in the mechanism of resistance to *M. perniciosa*.

However, it is interesting to note that the TgPR3 gene of cupuassu described by Silva et al. (2020) was found in this region. The cocoa orthologous gene was in the distal position of chromosome 6, around 30 kb from the marker 6M154484. This same marker is found in a more central position in cupuassu LG 6 due to chromosomal rearrangement at 14.55 cm which places it in the middle of the QTL region. By this set of correspondence, the TgPR3 gene can be positioned in the middle region of the WBD QTL (Fig. 5).

TgPR3 gene codes for a putative chitinase and is potentially related to the response to WBD in cupuassu (Silva et al. 2020). In addition, it is orthologous to the cacao gene Tc06v2_p000370 that is responsive to the attack of pathogens, being upregulated during cacao-*Phytophthora palmivora* interaction (Fister et al. 2016).

Given this configuration, the TgPR3 is a likely candidate gene for cupuassu resistance to *M. perniciosa*. The particularity of this analysis was that it was based on the same type of host-pathogen interactions, since the two clones used during differential expressions (Silva et al. 2020) were precisely the same two parents, 174 and 1074, of the progeny in our study. The strong degree of expression of those genes and the secretion of chitinase in resistant individuals affords a greater defense capacity and the possibility of being able to recognize pathogens, probably through the recognition of molecules derived from degradation of the pathogen's walls. The role of TgPR3, and its slightly lower resistant/susceptible differential expression, compared to TgPR5, can be placed upstream at the time of early pathogen detection. The constitutive expression of TgPR3 would make it possible, right from initial contact with the pathogen, to detect its presence and trigger the signaling pathway to induce a defense response through the secretion of new PR proteins (TgPR5 and TgPR8) that would block the pathogen. Such a mechanism would thus explain why we did not find the regions underlying the others PR genes.

Conclusions

With the combined use of high-throughput genotyping by sequencing (GBS) and a reference genome of a related species, *T. cacao*, we were able to develop a large SNP dataset for *T. grandiflorum* from a biparental population (174 resistant × 1074 susceptible clone to *M. perniciosa*). With this strategy, it was possible to construct the first high-density genetic map comprising 1438 SNP markers and 10 SSR, one of the prerequisites for producing a cupuassu reference genome,

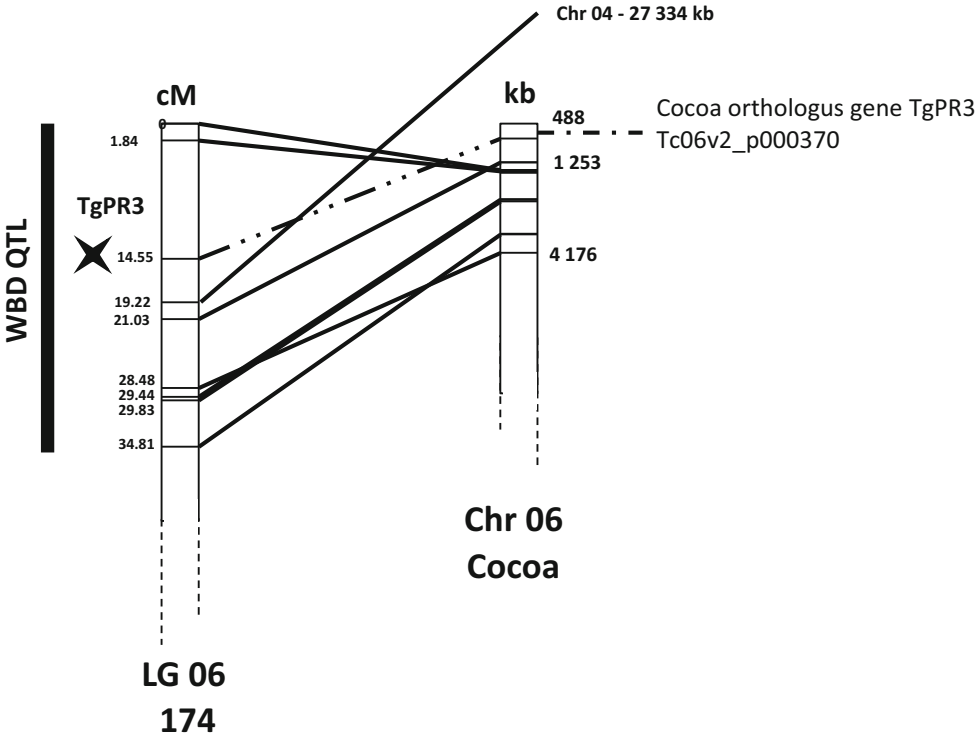

Fig. 5 QTL location in the cupuassu linkage group 6 over 9 markers and the position of the TgPR3 homolog in *T. cacao* genome. The homolog of the TgPR3 gene in cocoa is positioned 30 kb from the 6M488189 marker. This marker is located in the middle of the QTL zone detected on the WBD character at 14.55 cm

enabling the anchorage of a physical map. This tool also allowed to visualize the relations of synteny existing between the two species, *T. cacao* and *T. grandiflorum*. The genetic map also enables the detection and characterization of QTL associated with traits of interest for the genetic improvement of cupuassu. Here, a QTL of resistance to WBD was detected from observations on the crossing 174 × 1074. The analysis of this QTL, located in the distal region of chromosome 6 of parent 174, showed that the TgPR3 gene, potentially involved in the response to WBD in cupuassu (Silva et al. 2020), was co-localized with the QTL region. With these main results, it was thus possible to (i) provide the first *T. grandiflorum* high-density genetic map, (ii) specify a potential candidate gene involved in resistance to WBD in the interaction between *M. perniciosa* and *T. grandiflorum*, and (iii) contribute to developing marker-assisted selection strategies, enabling a pre-breeding at the young age of potentially resistant individuals. It would be interesting to check the possibility of cross-infection between *M. perniciosa*, from cocoa and cupuassu strains, and see if there is a parallel to this resistance in the cocoa tree. Finally, it is necessary to remain attentive to the evolution of this resistance and the interaction between *M. perniciosa* and *T. grandiflorum*. Indeed, it has been shown that after several years of interactions between *M. perniciosa* and *T. cacao*, resistance evolves from a monogenic to a polygenic expression.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11295-020-01479-3.

Acknowledgments We are grateful to the CEPLAC (*Comissão Executiva do Plano da Lavoura Cacaueira*), located in Marituba (Pará, Brazil), for installing the 174 × 1074 progeny in the field and making all the observations. This work was supported by the CIRAD-UMR AGAP HPC Data Center of the South Green Bioinformatics platform. We thank Fabienne Micheli (UESC/CIRAD) for its advice.

Funding We are grateful for the financial support from the tripartite project: Embrapa (Empresa Brasileira de Pesquisa Agropecuária), Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and Agropolis Foundation. We are grateful particularly for the financial contribution of Agropolis Foundation, which financed the whole of this study.

Compliance with ethical standards

Authorization No. AD93167 issued from a Sistema Nacional de Gestão do Patrimônio (SisGen) from Brasil, to useDNA extracted from T. grandiflorum leaves.

KeyGene N.V. CIRAD owns patents and patent applications protection its sequence-based genotyping technologies.

Data archiving statement Raw sequence data has been submitted to NCBI's Short Read Archive (SRA). Ref: PRJNA609621 https://www.ncbi.nlm.nih.gov/sra/PRJNA609621

The genetic map has been submitted to Tropgene db, database in genomic, genetic, and phenotypic information about tropical crops: https://tropgenedb.cirad.fr/tropgene/JSP/index.jsp

References

Aime MC, Phillips-Mora W (2005) The causal agents of witches broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of *Marasmiaceae*. Mycologia. 97(5):1012–1022. https://doi.org/10.3852/mycologia.97.5.1012

Albuquerque PSB, Silva AMC, Luz EDMN, Pires JL, Viera AMC, Demétrio CGB, Pascholatti SF, Figueira A (2009) Novel sources of witches' broom resistance (causal agent *Moniliopthora perniciosa*) from natural populations of Theobroma cacao from the Brazilian Amazon. Euphitica. 97:1012–1022. https://doi.org/10.3852/mycologia.97.5.1012

Allegre M, Argout X, Boccara M, Fouet O, Roguet Y, Berard A, Thevenin JM, Chauveau A, Rivallan R, Clement D (2012) Discovery and mapping of a new expressed sequence tag-single nucleotide polymorphism and simple sequence repeat panel for large-scale genetic studies and breeding of Theobroma cacao L. DNA Res 19:23–35

Alves RM, Resende MDV (2008) Avaliação genética de individuos e progênies de cupuaçuzeiro no estado do Pará e estimativas de parâmetros genéticos. Rev Bras Frutic 30:696–701. https://doi.org/10.1590/S0100-29452008000300023

Alves RM, Sebbenn AM, Artero AS, Figueira A (2006) Microsatellite loci transferability from *Theobroma cacao* to *Theobroma grandiflorum*. Mol Ecol Notes 6:583–586. https://doi.org/10.1111/j.1471-8286.2006.01496.x

Alves RM, Artero AS, Sebbenn AM, Clement C, Figueira A (2007) High levels of genetic divergence and inbreeding in populations of cupuassu, *Theobroma grandiflorum*. Tree Genet Genomes 3:289–298. https://doi.org/10.1007/s11295-006-0066-9

Alves RM, Resende MDV, Bandeira BS, Pinheiro TM, Farias DCR (2010) Avaliação e seleção de progênies de cupuaçuzeiro (*Theobroma grandiflorum*), em Belém, Pará. Rev Bras Frutic 32:204–212. https://doi.org/10.1590/S0100-29452010005000010

Alves RM, Silva CRS, Silva MSC, Silva DCS, Sebbenn AM (2013) Diversidade genética em coleções amazônicas de germoplasma de cupuaçuzeiro [Theobroma grandiflorum (Willd. ex Spreng.) Schum.]. Rev Bras Frutic 35:818–828. https://doi.org/10.1590/S0100-29452013000300019

Alves RM, Silva DCS, Albuquerque PSB, Santos VS (2017) Phenotypic and genotypic characterization and compatibility among genotypes to select elite clones of cupuassu. Acta Amazon 47:175–184. https://doi.org/10.1590/1809-4392201602104

Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X et al (2008) Towards the understanding of the cocoa transcriptome: production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao* L. generated from various tissues and under various conditions. BMC Genomics 9(1):512. https://doi.org/10.1186/1471-2164-9-512

Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J et al (2011) The genome of *Theobroma cacao*. Nat Genet 43:101–108. https://doi.org/10.1038/ng.736

Argout X, Martin G, Droc G, Fouet O, Labadie K, Rivals E, Aury JM, Lanaud C (2017) The cacao Criollo genome v2.0/ an improved version of the genome for genetic and finctional genomic studies. BMC Genomics 18:730. https://doi.org/10.1186/s12864-017-4120-9

Barbosa CS, Fonseca RR, Batista TM, Barreto MA, Argolo CS, Carvalho MR, Amaral DOJ, Silva EMA, Arévalo-Gardini E, Hidalgo KS, Franco GR, Pirovani CP, Micheli F, Gramacho KP (2018) Genome sequence and effectorome of Moniliophthora perniciosa and Moniliophthora roreri subpopulations. BMC Genomics 19:509–509. https://doi.org/10.1186/s12864-018-4875-7

Brown JS, Schnell RJ II, Motamayor JC, Lopes U, Kuhn DN (2005) Resistance gene mapping for witches' broom disease in

Theobroma cacao L. in an F2 population using SSR markers and candidate genes. J Am Soc Hortic Sci 130(3):366–373

Chaintreuil C, Rivallan R, Bertioli DJ, Klopp C, Gouzy J, Courtois B, Leleux P, Martin G, Rami JC, Gully D, Parrinello H, Severac D, Patrel D, Fardoux J, William Ribiere W, Boursot M, Cartieaux F, Czernic P, Pascal Ratet P, Mournet P, Giraud E, Arrighi JF (2016) A gene-based map of the Nod factor-independent *Aeschynomene evenia* sheds new light on the evolution of nodulation and legume genomes. DNA Res 23(4):365–376. https://doi.org/10.1093/dnares/dsw020

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Clement CR (1999) 1492 and the loss of Amazonian crop genetic resources. I. The relation between domestication and human population decline. Econ Bot 53:188–202. https://doi.org/10.1007/BF02866498

Cormier F, Lawac F, Maledon E, Gravillon MC, Nudol E, Mournet P, Vignes H, Chaïr H, Arnau G (2019) A reference high-density genetic map of greater yam (Dioscorea alata L.). Theor Appl Genet 132(6):1733–1744. https://doi.org/10.1007/s00122-019-03311-6

Costa MC, Maia GA, Souza Filho MSM, Figueiredo RW, Nassu R, Monteiro JCS (2003) Conservação de polpa de cupuaçu [*Theobroma grandiflorum* (Willd. Ex Spreng.) Schum] por métodos combinados. Rev Bras Frutic 25:213–215. https://doi.org/10.1590/S0100-29452003000200007

Costa MP, Frasao BS, Rodrigues BL, Silva ACO, Conte-Junior CA (2016) Effect of different fat replacers on the physicochemical and instrumental analysis of low-fat cupuassu goat milk yogurts. J Dairy Res 83(4):493–496. https://doi.org/10.1017/S0022029916000674

Da Silva RA, Souza G, Lemos LSL, Lopes UV, Patrocino NGRB, Alves RM, Marcellino HL, Clément D, Micheli F, Gramacho KP (2017) Genome size cytogenetic data and transferability of EST-SSRs markers in wild and cultivated species of the genus *Theobroma* L. (Byttnerioideae, Malvaceae). PLoS One 12(2):e0170799. https://doi.org/10.1371/journal.pone.0170799

De Oliveira TB, Genovese MI (2013) Chemical composition of cupuassu (*Theobroma grandiflorum*) and cocoa (*Theobroma cacao*) liquors and their effects on streptozotocin-induced diabetic rats. Food Res Int 51(2):929–935. https://doi.org/10.1016/j.foodres.2013.02.019

Falcão LL, Werneck JOS, Togawa RC, Costa MMdC, Grynberg P, Silva Junior OBd, Alves RM, Albuquerque PSB, Marcellino LH (2018). Analyses of cupuassu (*Theobroma grandiflorum*) transcriptome during interaction with *Moniliophthora perniciosa*, the causal agent of witches' broom disease. In: Cpatu S-E (Ed) 7th brazilian biotechnology congress and 2nd biotechnology ibero-american congress, Brasília. DF, Brazil

Ferraz Dos Santos L, Moreira Fregapani R, Falcão LL, Togawa RC, Costa MMC, Lopes UV, Peres Gramacho K, Alves RM, Micheli F, Marcellino LH (2016a) First microsatellite markers developed from cupuassu ESTs: application in diversity analysis and cross-species transferability to cacao. PLoS One 11:e0151074–e0151074. https://doi.org/10.1371/journal.pone.0151074

Ferraz dos Santos L, Santana Silva RJ, Oliveira Jordão do Amaral D, Barbosa de Paula MF, Ludke Falcão L, Legavre T, Alves RM, Marcellino LH, Micheli F (2016b) Selection of reference genes for expression study in pulp and seeds of theobroma grandiflorum (willd. Ex spreng.). Schum. PLoS One 11:e0160646. https://doi.org/10.1371/journal.pone.0160646

Fister AS, Mejia LC, Zhang Y, Herre EA, Maximova SN, Guiltinan MJ (2016) Theobroma cacao L. pathogenesis-related gene tandem array members show diverse expression dynamics in response to pathogen colonization. BMC Genome 17:363. https://doi.org/10.1186/s12864-016-2693-3

Frias G, Purdy L, Schmidt R (1995) An inoculation method for evaluating resistance of cacao to *Crinipellis perniciosa*. Plant Dis 79:787–791

Garsmeur O, Droc G, Antonise R, Grimwood J, Potier B, Aitken KS, Jenkins J, Martin G, Charron C, Hervouet C, Costet L, Yahiaoui N, Healey A, Sims D, Cherukuri Y, Sreedasyam A, Kilian A, Chan AP, Van Sluys MA, Swaminathan K, Town C, Berges H, Simmons B, Glaszmann JC, Van der Vossen E, Henry R, Schmutz J, D'Hont A (2018) A mosaic monoploid reference sequence for the highly complex genome of sugarcane. Nat Commun 9:2638, 10. https://doi.org/10.1038/s41467-018-05051-5

Gesteira AS, Micheli F, Carels N, Da Silva A, Gramacho K, Schuster I et al (2007) Comparative analysis of expressed genes from cacao meristems infected by *Moniliophthora perniciosa*. Ann Bot 100(1):129–140. https://doi.org/10.1093/aob/mcm092

Gilabert-Escrivá MV, Gonçalves LAG, Silva CRS, Figueira A (2002) Fatty acid and triacylglycerol composition and thermal behaviour of fats from seeds of Brazilian Amazonian *Theobroma*. J Sci Food Agric 82:1421–1431. https://doi.org/10.1002/jsfa.1107

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS One 9(2):e90346. https://doi.org/10.1371/journal.pone.0090346

Gramacho KP, Luz EDMN, Silva FS, Lopes UV, Pires JL, Santos-Filho LP (2016) Pathogenic variability of *Moniliophthora perniciosa* in three agroecological zones of the cacao region of Bahia, Brazil. Crop Breed Appl Biotechnol 16:7–13. https://doi.org/10.1590/1984-70332016v16n1a2

Gu Z (2014) Circlize implements and enhances circular visualization in R. Bioinformatics 30:2811–2812. https://doi.org/10.1093/bioinformatics/btu393

Kuhn DN, Figueira A, Lopes U, Motamayor JC, Meerow AW, Cariaga K, Freeman B, Livingstone DS III, Schnell RJ (2010) Evaluating T*heobroma grandiflorum* for comparative genomic studies with *Theobroma cacao*. Tree Genet Genomes 6(5):783–792. https://doi.org/10.1007/s11295-010-0291-0

Leal GA, Gomes LH, Albuquerque PSB, Tavares FCA, Figueira A (2010) Seraching for Moniliophthora perniciosa pathogenicity genes. Fugal Biol 114(10):842–854. https://doi.org/10.1016/j.funbio.2010.07.009

Lopes UV, Wilson RM, Pires JL, Clement D, Yamada MM, Gramacho KP (2011) Cacao breeding in Bahia, Brazil - strategies and results Crop Breed. Appl Biotechnol 11(spe) Viçosa June 2011. https://doi.org/10.1590/S1984-70332011000500011

Maia MCC, Resende MDV, Oliveira LC, Alves RM, Silva Filho JL, Rocha MM, Cavalcante JJV, Roncatto G (2011) Análise genética de famílias de meios-irmãos de cupuaçuzeiro. Pesquisa Florestal Brasileira 31:123–130. https://doi.org/10.4336/2011.pfb.31.66.123

Montego JMC et al (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into witches' broom disease of cacao. BMC Genomics 9:548. https://doi.org/10.1186/1471-2164-9-548

Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D, Cornejo O et al (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. https://doi.org/10.1186/gb-2013-14-6-r53

Nascimento AS, Alves RM, Silva CRS, Albuquerque PSB, Lemos LSL, Gramacho KP, Clément D (2019) Transferability and mapping of microsatellites between *Theobroma cacao* L and Theobroma grandiflorum (Willd. Ex. Spreng) Schum. Agrotrópica 31(2):85–102. https://doi.org/10.21757/0103-3816.2019v31n2p85-102

Neto AAP, Laranjeira D, Pires LEDMN (2018) Selection of cocoa progenies for resistance to witches' broom. Trop Plant Pathol 43:381–388. https://doi.org/10.1007/s40858-017-0208-7

Ollitrault P, Terol J, Chen C, Federici TC, Lotfy S, Hippolyte I, Ollitrault F, Bérard A, Chauveau A, Cuenca J, Costantino G, Kacar Y, Mu L, Garcia-Lor A, Froelicher Y, Aleza P, Boland A, Billot C, Navarro L, Luro F, Roose ML, Gmitter FG, Talon M, Brunel D (2012) A reference genetic map of C. Clementina hort. Ex Tan.; citrus

evolution inferences from comparative mapping. BMC Genomics 13:593. https://doi.org/10.1186/1471-2164-13-593

Pires JL, De Melo GP, Yamada MM, Gramacho KP (2012) Association among sources of resistance to witches' broom disease for the increment of the level and durability of the character. Agrotropica 24: 27–30. https://doi.org/10.21757/0103-3816.2012v24n1p27-30

Pugliese AG, Tomas-Barberan FA, Truchado P, Genovese MI (2013) Flavonoids, proanthocyanidins, vitamin C, and antioxidant activity of *Theobroma grandiflorum* (Cupuassu) pulp and seeds. J Agric Food Chem 61(11):2720–2728. https://doi.org/10.1021/jf304349u

Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high-density linkage map of Theobroma cacao L. Theor Appl Genet 101(5/6):948–955. https://doi.org/10.1007/s001220051566

Royaert S, Jansen J, da Silva DV, de Jesus Branco SM, Livingstone DS III, Mustiga G, Marelli JP, Ioná Araújo IS, Xorrêa 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

Salgon S, Jourda C, Sauvage C, Daunay MC, Reynaud B, Wicker E, Dintinger J (2017) Eggplant resistance to the Ralstonia solanacearum species complex involves both broad-spectrum and strain-specific quantitative trait loci. Front Plant Sci 8:828, 21. https://doi.org/10.3389/fpls.2017.00828

Santos FFJ, Lopes UV, Pires JL, Melo GR, Gramacho KP, Clément D (2014) QTL detection under natural infection of *Moniliophthora perniciosa* in a cacao F2 progeny with Scavina-6 descendants. Agrotrópica 26(1):65–72 Centro de Pesquisas do Cacau, Ilhéus, Bahia, Brasil

Silva RJS, Alves RM, Gramacho KP, Marcellino LH, Micheli F (2020) Involvement of structurally distinct cupuassu chitinases and osmotin in plant resistance to the fungus *Moniliophthora perniciosa*. Plant Physiol Biochem 148:142–151. https://doi.org/10.1016/j.plaphy.2020.01.009

Sveinsson S, Gill N, Kane NC, Cronk Q (2013) Transposon fingerprinting using low coverage whole genome shotgun sequencing in Cacao (*Theobroma cacao* L.) and related species. BMC Genomics 14:502. https://doi.org/10.1186/1471-2164-14-502

Teixeira PJPL, Thomazella DPT, Reis O, do Prado PFV, do Rio MCS, Fiorin GL, José J, Costa GGL, Negri VA, Mondego JMC, Mieczkowski P, Pereira GAG (2014) High-resolution transcript profiling of the atypical biotrophic interaction between *Theobroma cacao* and the fungal pathogen *Moniliophthora perniciosa*. Plant Cell 26(11):4245–4269. https://doi.org/10.1105/tpc.114.130807

Van Ooijen JW (2006) JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Kvazma B.C., Wageningen