GMR
The Original by FUNPEC-RP

Comparing SNP-based genetic linkage and physical maps in guava (Psidium guajava)

F.L.B. Medeiros[1], C.A.F. Santos[2] and A.E.S. Costa[1]

[1] Universidade Federal Rural de Pernambuco, Departamento de Agronomia, Dois Irmãos, Recife, PE, Brasil
[2] Embrapa Semiárido, Petrolina, PE, Brasil

Corresponding author: C.A.F. Santos
E-mail: carlos-fernandes.santos@embrapa.br

Genet. Mol. Res. 21 (2): gmr19033
Received February 10, 2022
Accepted March 23, 2022
Published May 25, 2022
DOI http://dx.doi.org/10.4238/gmr19033

ABSTRACT. Genome mapping is a simplified representation of molecular markers or nucleotide sequences in chromosomes; developing accurate and dense maps is crucial for marker-assisted selection. We developed and compared genetic linkage maps obtained using JoinMap 4.0 and GACD with a physical map obtained using BLAST analysis based on *Eucalyptus* SNPs transferred to guava, *Psidium guajava* (Myrtaceae) - to serve as a reference for trait mapping in this crop. Genotyping was conducted on 112 individuals from an experimental cross between a well-known commercial cultivar and an exotic genotype (Pedro Sato × Purple guava), using the Euchip60K SNP chip, version 2.0 (72,202 SNPs); 79% of the SNPs were monomorphic. After data filtering, 1120 markers were used for map construction. The JoinMap 4.0 linkage map had 203 markers, spanning 1405.2 cM, with an average marker distance of 7.7 cM. The GACD linkage map had 186 markers and spanned 1392.7 cM, with an average marker distance of 8.8 cM. JoinMap and GACD disagreed on the estimated distances and SNP ordering. GACD showed a greater limitation than JoinMap 4.0 as it ordered markers according to their parental origin. The physical map developed using BLAST consisted of 694 hits (e-values from $8xE^{-10}$ to $1.15xE^{-26}$), spanning 434.88 Mb, with an average marker interval of 0.62 Mb. Both linkage maps showed linkage groups with segments from

several chromosomes compared to the physical map, indicating limitations. These results highlight the effectiveness of physical mapping through BLAST to overcome linkage mapping limitations, such as in marker grouping and ordering. The physical maps proposed here can serve as a reference for mapping and QTL estimates in guava.

Key words: *Eucalyptus*; Guava; SNPs transferability, BLAST

INTRODUCTION

Guava (*Psidium guajava* - Myrtaceae) is a major crop in the Brazilian fruit sector and has its probable center of origin in South America (Risterucci et al., 2005). Both fresh and industrialized guava consumption are economically important. The fruits have high nutritional value due to ascorbic acid and lycopene content (Masud Parvez et al., 2018). Brazil is one of the largest *P. guajava* producers in the world, and Paluma and Pedro Sato are the most widely planted cultivars. Feng et al. (2020) have recently sequenced the *P. guajava* genome (cultivar New Age) and assembled 443.8 Mb into 11 chromosomes (n=11), representing 95.7% of the genome, placing guava at the same level as other important crops, e.g., eucalyptus (Myburg et al., 2014).

Genome mapping is a graphical representation of marker or nucleotide sequence position in the chromosomes (Lander et al., 2001). From this perspective, two approaches are available for mapping purposes: physical mapping and linkage mapping (genetic linkage mapping). Physical maps differ from genetic maps for reflecting the actual distance between nucleotides, whereas genetic maps are based on chromosome recombination rates (Bueno, 2009).

Linkage mapping reports are available for various Myrtaceae species, e.g., *Eucalyptus* spp. (Sumathi et al., 2018), *Leptospermum scoparium* (Chagné et al., 2019), and *P. guajava* (Rodríguez et al., 2007; Lepitre et al., 2010; Ritter et al., 2010a; Padmakar et al., 2015). Paiva et al. (2011) reported the construction of a bacterial artificial chromosome (BAC) for physical mapping in *E. grandis* (BRASUZ1 clone) to assemble the reference genome of the species. However, no reports of physical mapping or BAC libraries were found for *P. guajava*.

According to Hohmann et al. (1994), comparing physical and linkage maps is a way to combine genetic mapping and molecular information of different species. From this perspective, DeWan et al. (2002) reported inconsistencies between physical and linkage maps of panels 9 and 10 of the human genome using data from the Human Genome Project – Santa Cruz and Celera databases. Moreover, comparisons between an SNP-based linkage map in *Capsicum baccatum* and the physical map of *C. annuum* also revealed inconsistencies as *C. baccatum* had chromosomal segments from Chr. 3, Chr. 9, and Chr. 5 of *C. annuum* (Lee et al., 2016).

The Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) can be used to assemble physical maps or for partial marker allocation in chromosomes according to the similarities observed (hits), as demonstrated by Gao et al. (2018) when using BLAST to physically locate scaffolds and SNP markers of *Pleurotus tuoliensis* by adopting an e-value of $1xE^{-3}$. In another study, Blair et al. (2018) reported multiple BLAST hits for 66 of

768 SNPs, indicating paralogue sequences that could cause inconsistencies in the grouping and ordering of linkage groups. However, reports on the development of an SNP-based physical map for *P. guajava* were not found, making comparisons with linkage maps impossible.

From this perspective, this study aimed to develop and compare linkage maps, obtained with JoinMap 4.0 and GACD programs, with a physical map obtained through BLAST analysis of *Eucalyptus* SNPs against the guava genome, to serve as a reference for mapping qualitative and quantitative traits in guava.

MATERIAL AND METHODS

Plant material, DNA extraction and quantification

The plant material was obtained from a experimental cross between the guava cultivars 'Pedro Sato' × 'Purple guava' (PSR) from the Active Germplasm Bank of Embrapa Semi-arid (Embrapa Semiárido) - Petrolina, PE, Brazil. The cultivar 'Pedro Sato' has red fruit pulp, whereas the cultivar 'Purple guava' has purplish leaves, fruit, stem, and pulp. Besides their phenotypical differences, Pedro Sato is a well-established economic important cultivar in Brazil that has Purple guava's disease resistance and could generate potential materials for guava breeding. Guava progenies were grafted on BRS Guaraçá rootstock four months after transplanting; this rootstock is resistant to the root-knot nematode (*Meloidogyne enterolobii*).

Healthy leaves from 112 individuals were sampled for DNA extraction according to the protocol proposed by Inglis et al. (2018) in which 10-150g of fresh leaf tissue were carefully macerated using liquid nitrogen. Sampling was done three months after grafting. DNA quantification was performed by spectrophotometry using a micro-drop plate at an absorbance of 260 nm. After quantification, the samples were diluted to 20 ng/ μL.

SNP genotyping and development of genetic linkage maps

Genotyping was conducted using the Euchip60K DNA marker chip, version 2.0, with 72202 SNPs (Silva-Junior et al., 2015). Genotyping services were performed by GeneSeek (Lincoln, NE, USA). Monomorphic markers and those with unexpected segregation or missing data in >20% progenies were eliminated. Markers were identified with "A" and "C" at the end of the SNP's name to identify the parental origin. In the case of codominance, the markers were identified with "H".

Linkage analysis performed with JoinMap 4.0 (Van Ooijen, 2006) followed this script: 1) first, markers with chi-square >20 for the ll, lm, nn, and np phases, and chi-square >70 for the hh, hk, and hk phases were eliminated from the analysis; 2) then, the data were coded for cross-pollination (CP); 3) next, grouping was performed based on LOD scores from 5.0 to 12; 4) then, the markers were ordered in groups using maximum likelihood; 5) next, markers with N.N. stress (cM) > 4.0 were eliminated; 6) finally, the values obtained from recombination frequencies were converted into map distances using Kosambi's function (1943).

Linkage analysis with GACD (Zhang et al., 2015) followed this script: 1) first, grouping was performed based on a LOD score of 8.0; 2) then, marker ordering used the "k-

optimality" criterion based on the LOD score, 3-OptMAP, and the N.N. (nearest neighbor) algorithm; 3) the rippling criterion by LOD was "windows size 8"; 4) finally, recombination frequency values were converted into map distances using Kosambi's function (1943).

In silico SNP analysis and physical mapping in *P. guajava*

In silico analysis was performed via BLAST by aligning *Eucalyptus* SNPs (query) against the guava reference genome (subject) (Feng et al., 2020) using data from the National Center for Biotechnology Information (NCBI). Each SNP was approximately 70 nucleotides long.

The BLAST tool was then applied (optimization: BLASTN, expected threshold = 1) to estimate SNP positions. Hits between *Eucalyptus* SNPs and the guava genome were analyzed in an Excel spreadsheet. Markers were ordered according to the guava chromosomes. A cut-off e-value $<E^{-10}$ was adopted to exclude markers. Markers with close e-values were kept in the map and identified by "*" if in the same chromosome and by "#" if in different chromosomes. Finally, physical maps were built with the mapping software MapGene2Chrom (Jiangtao et al., 2015).

RESULTS

After data filtering, 1120 SNPs (1.6%) of a total 72202 SNPs of the EUChip60K (Silva-Junior et al., 2015) were used for linkage and physical mapping. In the Pedro Sato × Purple guava experimental cross, 79% of the SNPs were monomorphic, while the others showed unexpected Mendelian segregation rates or missing data in >20% of the progenies.

Genetic linkage mapping

The JoinMap 4.0 linkage map showed 203 SNP markers in 11 guava linkage groups, spanning 1405.2 cM. The average marker distance was 7.7 cM. Linkage group length ranged from 49.3 (LG9) to 196.8 cM (LG7) and from 11 to 26 markers per group. The distance between markers ranged from 0.1 to 28.4 cM (Figure 1 and supplementary data).

The second linkage map, generated using GACD, was 1392.7 cM long, with 186 markers spanning all eleven linkage groups. The average marker distance was 8.8 cM. Linkage groups 5 and 2 were the densest, whereas the least dense was group 11. LG1 was the shortest group (49.86 cM), whereas LG5 was the longest (180.80 cM). The distance between markers ranged from 0.94 to 37.93 cM (Figure 2 and supplementary data).

A mix of coupling and repulsion phases was observed in the JoinMap 4.0 linkage map, identified by the "C" and "A" markers from different parents. On the other hand, this mix was not observed in the GACD linkage map, and markers were ordered according to the parental origin (Figures 1 and 2 and supplementary data). There were 91 type "A" parental dominant markers, 75 type "C" dominant markers, and 37 co-dominants (H) in the JoinMap 4.0 linkage map. On the other hand, there were 88 "A" markers, 81 "C" markers, and 17 "H" markers in the GACD map (Figures 1 and 2 and supplementary data). With regard to the 37 "H" markers (JoinMap 4.0), 25 did not match with codominant markers mapped by GACD.

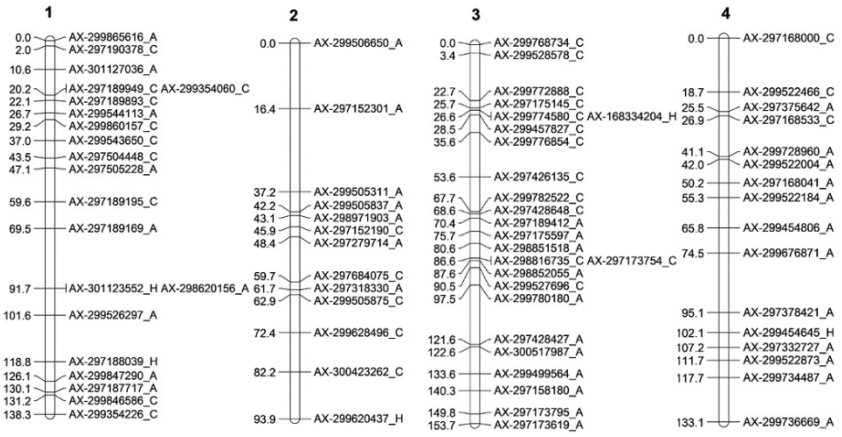

Figure 1 Guava genetic linkage groups 1, 2, 3 and 4 from the Pedro Sato × Purple guava (PSR) cross obtained with JoinMap 4.0 (Van Ooijen 2006).

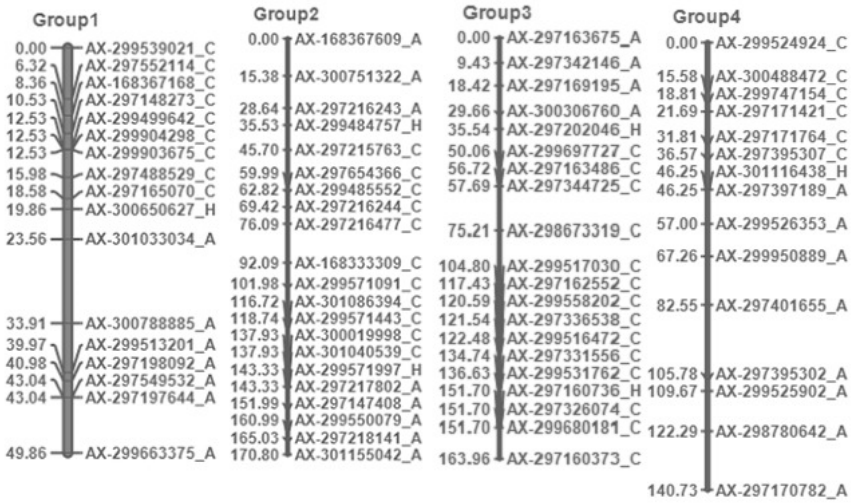

Figure 2 Guava genetic linkage groups 1, 2, 3 and 4 from the Pedro Sato × Purple guava (PSR) cross obtained with GACD (Zhang et al., 2015).

There were common markers in all linkage groups, except JoinMap LG6 that did not have correspondence with GACD linkage groups (Table 1). Among the correspondent groups, 85.7% of markers mapped to LG1 (JoinMap 4.0) were also mapped in LG9 (GACD), 92.3% of LG2 markers (JoinMap 4.0) were mapped in LG8 (GACD) (Figures 1 and 2 and supplementary data). Groups 3, 4 and 11 (JoinMap 4.0) had 100% of their markers mapped in groups 5, 6 and 10 (GACD), respectively. LG5 (JoinMap 4.0) presented 71.4% of its markers mapped in LG11 (GACD). LG7 (JoinMap 4.0) had 80% of markers mapped in LG3 (GACD) (Figures S1 and S2 - supplementary data). In LG8 (JoinMap 4.0) 95.5% of the markers were equal to LG2 (GACD), as well as 94.1% of LG9 corresponded to LG1 (GACD). Lastly, 92.3% of markers from LG10 (JoinMap 4.0) were equal to LG7 (GACD) (Figures 1 and 2 and supplementary data).

Table 1. Linkage group length (LGL), marker number/linkage group and number of codominant markers (H) in the genetic linkage maps from the cross Pedro Sato × Purple guava cultivars using the softwares JoinMap 4.0 × GACD.

JoinMap 4.0				GACD			
LG	**LGL (cM)**	**Marker number**	**Markers (H)**	**LGL (cM)**	**Marker number**	**Markers (H)**	**LG**
LG 1	138.3	21	2	49.86	17	1	9
LG 2	93.9	13	1	170.8	21	2	8
LG 3	153.7	25	1	163.9	20	2	5
LG 4	133.1	16	1	140.73	15	1	6
LG 5	147.3	14	1	180.8	29	3	11
LG 6	180.6	26	23	141.2	16	1	No
LG 7	196.8	25	3	93.9	13	1	3
LG 8	174.9	22	2	133.17	14	1	2
LG 9	49.3	17	1	180.11	19	2	1
LG 10	84.4	13	1	51.65	12	1	7
LG 11	52.9	11	1	86.6	10	1	10
Total	1405.2	203	37	1392.7	186	17	

cM: centimorgan.

Different orderings were observed in almost all linkage groups built with JoinMap 4.0 and GACD such as: LG1 x LG9, LG2 x LG8, LG3 x LG5, LG4 x LG6, LG7 x LG3, LG8 x LG2, LG9 x LG1, LG10 x LG7 and LG11 x LG10 (Figures 1 and 2 and supplementary data).

Genetic linkage maps also presented a few similar orderings like LG4 (JoinMap 4.0) in which six markers were ordered in the same as LG6 (GACD). This situation was observed in all the other groups, in different proportions, such as in LG3 x LG5 and LG8 x LG2 that presented three and two markers ordered similarly, respectively. On the other hand, LG9 x LG1 did not present coincidences in marker ordering. Linkage groups 5 x 11 and 11 x 10 showed the highest coincidence in SNP ordering between maps, as linkage group 5 (JoinMap 4.0) was almost identical to LG11 (GACD), with the exception of four markers; and linkage group 11 had 54% of its markers ordered similarly to LG10 (Figures 1 and 2 and supplementary data).

In LG4 x LG6 marker positions differ almost entirely, with the exception of markers AX-297332727, positioned at 107.2 cM (LG4) and 110.97 cM (LG6) and AX-299454645 positioned at 102.2 cM (LG4) and 115.97 cM (LG6). In LG5 x LG11 marker AX-297213963 was mapped in the same position in both linkage maps. The greatest difference observed between these groups was marker AX-297182690 position: 102.7 cM (LG5) and 86.87 cM (LG11), however the remaining markers in this group were close in relative distances, with differences ranging from 3.51 cM (AX-299521130) to 15 cM (AX-168367414 and AX-297182690).

Regarding LG11 x LG10 they presented nearly identical marker positions. Marker AX-301123916 was the first in both linkage groups, the greatest observed relative distance was between marker AX-299565823, mapped at 52.9 cM (LG11) and 44.96 cM (LG10) (Figures S1 and S2 – supplementary data), the remaining markers presented even closer relative distances, ranging from 1.15 cM (AX-299563711) to 7.94 cM (AX-297616735) (Figures 1 and 2 and supplementary data).

Physical mapping of SNPs *Eucalyptus* in *P. guajava*

In silico analysis of *Eucalyptus* SNPs generated 694 hits from the 1120 initially filtered markers for linkage analysis. These alignments had e-values ranging from $8xE^{-10}$ to $1.15xE^{-26}$ and were ordered to build the first guava physical map (Figure 3 and supplementary data).

Chromosome 3 was the longest in length (50.58 Mb) among the eleven guava chromosomes, meanwhile chromosome 9 was the shortest with 32.33 Mb (Table 2). The aligned SNPs covered almost the entire guava genome (more than 90% of the reported chromosome length), ranging from 31.4 Mb (chromosome 9) to 49.27 Mb (chromosome 3) and covering an average of 39.5 Mb/chromosome (Table 2). Marker number per linkage group ranged from 40 (chromosome 8) to 88 (chromosomes 5 and 7), with an average of 63 SNPs per chromosome (Table 2). Average marker distance was 0.62 Mb, ranging from 0.01 Mb to 7.36 Mb.

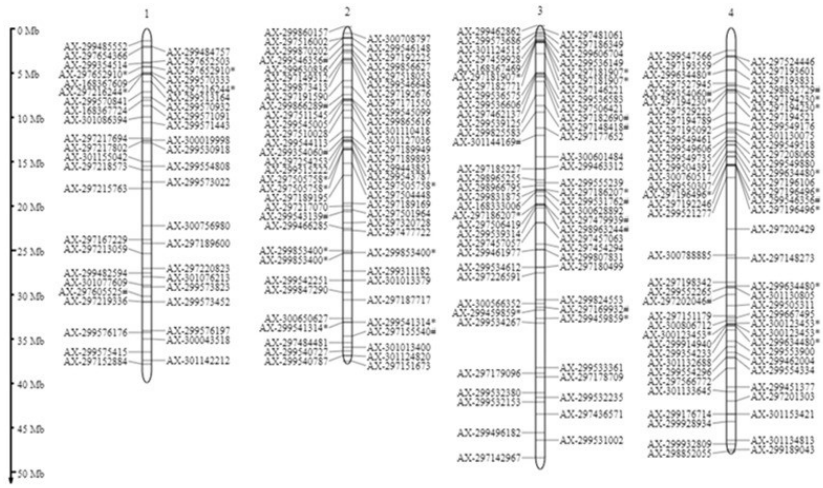

Figure 3. Physical map of chromosomes 1, 2, 3, 4, 5 and 6 of Pedro Sato × Purple guava cross (PSR), obtained by BLAST using *Eucalyptus* SNPs against the whole guava genome with MapGene2Chrom (Jiangtao et al., 2015). Markers identified with "*" are repeated in the same chromosome, whereas markers identified with "#" are repeated in more than one chromosome.

Table 2. *Psidium guajava* chromosome length reported by Feng et al. (2020), physical map coverage and *Eucalyptus* SNP markers/chromosome of *Psidium guajava*.

Chromosome	Total chromosome length (Mb)	Physical map coverage (Mb)	SNP markers/chromosome
1	40.37	38.21	45
2	38.47	38.13	62
3	50.58	49.27	65
4	48.29	48.03	69
5	44.77	44.46	88
6	42.82	42.80	63
7	35.36	35.59	88
8	33.38	32.96	40
9	32.33	31.4	50
10	37.88	37.78	71
11	37.02	36.25	53
Average	40.12	39.53	63
Total	441.27	434.88	694

Mb: megabase.

Comparison between linkage and physical maps

The physical map was denser than both linkage maps (Table 3). The Joinmap 4.0 linkage map showed an average of 18 SNPs per linkage group, of which group LG6 was the

densest, followed by LG3, LG7, and LG11, the least dense (Table 3). Linkage map GACD showed an average of 16 SNPs per linkage group and ranged from 10 (LG11) to 29 markers per group (LG5). The physical map showed an average of 63 SNPs per chromosome, ranging from 40 (chromosome 8) to 88 markers (chromosomes 5 and 7) (Table 3).

Table 3. SNPs per chromosome/linkage group in different genetic and physical maps of *Psidium guajava*.

Chromosome/linkage group	SNPs per chromosome/linkage group		
	JoinMap 4.0	GACD	Physical
1	21	17	45
2	13	21	62
3	25	20	65
4	16	15	69
5	14	29	88
6	26	16	63
7	25	13	88
8	22	14	40
9	17	19	50
10	13	12	71
11	11	10	53
SNP average	18	16	63
Total map length	1405.2 cM	1392.7 cM	434.88 Mb

cM: centimorgan.

The linkage map from JoinMap 4.0 had 122 markers in common with the physical map. Some of these linkage groups had SNP sequences distributed in different guava chromosomes, e.g., LG6, which had SNP sequences in eight chromosomes: 1, 2, 3, 5, 6, 7, 9, and 11 (Table 4). The GACD linkage map had 117 markers in common with the physical map, and a similar situation was observed as LG1 had SNP sequences distributed in chromosomes 2, 4, 6, 9, and 10 (Table 5).

Table 4. Comparison between the JoinMap 4.0 linkage map × physical map of *Psidium guajava*.

Linkage group	Linkage group length (cM) - JoinMap	Corresponding chromosomes in the physical map
1	138.3	Chr. 2
2	93.9	Chr. 4, Chr. 5
3	153.7	Chr. 4, Chr. 7
4	133.1	Chr. 9
5	147,3	Chr. 3, Chr. 4, Chr. 9, Chr. 11
6	180.6	Chr. 1, Chr. 2, Chr. 3, Chr. 5, Chr. 6, Chr. 7, Chr. 9, Chr. 11
7	196.8	Chr. 2, Chr. 3, Chr. 4, Chr. 5, Chr. 6, Chr. 8
8	174.9	Chr. 1, Chr. 9, Chr. 10
9	49.3	Chr. 2, Chr. 4, Chr. 6, Chr. 9, Chr. 10
10	84.4	Chr.1 Chr. 3, Chr. 6, Chr. 11
11	52.9	Chr. 11

cM: centimorgan.

Three linkage groups in JoinMap 4.0 (LG1, LG4, and LG11) had their markers matched in only one chromosome of the physical map (Table 4). The comparison between the JoinMap 4.0 linkage map and the physical map reveals a 57% correspondence between LG1 and chromosome 2 and similar ordering between the estimates obtained by

recombination frequencies and the actual position in the genome, except for markers AX-299544113 (12.5 Mb) and AX-299860157 (0.74 Mb). LG4 x chromosome 9 showed 62.5% of marker correspondence. However, the order was not similar. Nevertheless, LG11 x chromosome 11 showed 72.7% of marker correspondence, although with changed ordering.

Table 5. Comparison between the GACD linkage map × physical map of *Psidium guajava*.

Linkage group	Linkage group length (cM) - GACD	Corresponding chromosomes in the physical map
1	49.86	Chr. 2, Chr.4, Chr.6, Chr.9, Chr. 10
2	170.80	Chr. 1, Chr. 9, Chr. 10
3	163.9	Chr. 3, Chr. 4, Chr. 6, Chr. 8
4	140.73	Chr. 7
5	180.80	Chr. 4, Chr. 6, Chr. 7
6	141.20	Chr. 9
7	93.59	Chr. 1, Chr. 3, Chr. 6, Chr. 11
8	133.17	Chr. 4, Chr. 5, Chr. 6
9	180.11	Chr. 2
10	51.65	Chr. 8, Chr. 11
11	86.87	Chr. 3, Chr. 4, Chr. 9, Chr. 11

cM: centimorgan.

The comparison between the GACD linkage map and the physical map reveals that all linkage groups showed SNP sequences in more than one chromosome, except LG4, LG6, and LG9 (Table 5). The SNP order was not maintained between LG4 x chromosome 7 (53.3% marker correspondence), as well as between LG 6 x chromosome 9 (62.5%), except for markers AX-299454645, AX-297168533, AX-299522466, and AX-29716800, sequentially ordered in both maps (Figure 2 and S4 - supplementary data). The position of marker AX-299454645 was 115.97 cM (LG6), or 25.37 Mb in the physical map. Lastly, the LG9 x chromosome 2 correspondence level was 52.6%, with changes in SNP ordering.

DISCUSSION

This is the first report of the construction of an SNP-based linkage and physical maps for *P. guajava* by associating recombination frequency distances with the actual position in the guava genome and comparing the results obtained using two different mapping algorithms. Map comparisons aimed to evaluate the accuracy of traditional linkage mapping in guava, based on the real distances portrayed by the physical map. Precise mapping for marker-assisted selection depends on reliable estimations among marker and traits. On the other hand, no comparisons between JoinMap and GACD were found in the specialized literature.

The LOD scores in the JoinMap 4.0 and GACD maps ranged from 5 to 12 and 8, respectively, values equal to or greater than those found in the literature for guava linkage maps (Lepitre et al., 2010; Padmakar et al., 2015).

The comparison between the JoinMap 4.0 and GACD linkage maps reveals similar numbers of total linked SNPs (203 and 186, respectively) and markers per linkage group (18 SNPs/LG and 16 SNPs/LG, respectively). However, the order and distance between markers diverged significantly, with few exceptions (e.g., LG5 x LG11). Marker ordering is similar to the 'traveling salesman' problem, in which the best route must be chosen from a

set of m!/2 possibilities (Van Ooijen and Jansen, 2013). Using the maximum likelihood function, the software JoinMap 4.0 was more efficient in SNP ordering in *P. guajava* than GACD as the latter ordered markers according to their parental origin. However, this does not have a biological foundation since recombination, resulting from meiosis, is random process and unlikely to generate markers linked and positioned according to one of the parents. Furthermore, the GACD map had fewer codominant markers (17) than JoinMap 4.0 (37), and such markers are regarded as more informative for linkage mapping (Wu et al., 2002). Moreover, parental map integration depends on codominant markers to join the coupling and repulsion phases.

Guava linkage maps previously reported by Lepitre et al. (2010) and Padmakar et al. (2015), based on SSR and SRAP, were estimated to span 2179 cM and 2551.3 cM, respectively, longer than the maps reported in the present study (1405.2 cM and 1392.7 cM). The SNP-based linkage map for *L. scoparium*, a species of the Myrtaceae family, resulted in two parental maps, with 1242.8 and 1616.2 cM (Chagné et al., 2019), similar to the present study.

The construction of physical maps using SNP sequence alignments could be an alternative to estimate the actual marker position in species whose whole genome sequences are available. BLAST was efficient for this purpose in the present study, with e-values ranging from $8xE^{-10}$ to $1.15xE^{-26}$. Gao et al. (2018) physically located SNPs approximately 20 nucleotides long in scaffolds of *P. tuolensis* using a cut-off e-value $<E^{-3}$. In the present study, the SNP sequences were 70 nucleotides long, and a cut-off e-value $< E^{-10}$ was adopted, which was effective in identifying biologically significant alignments for constructing physical maps.

The comparison between the JoinMap 4.0 linkage map and the physical map revealed 60.1% (122 SNPs) of common markers and a significant correspondence between LG11 and chromosome 11 and between LG1 and chromosome 2. Moreover, the GACD map had 117 markers in common with the physical map (62.9%) and no correspondence between LGs and chromosomes. Grattapaglia et al. (2015) used BLAST to identify consistencies between linkage and physical maps of *E. grandis* (available at Phytozome) and observed that 94.7% of the EST-SSR were associated with the physical map. Due to marker specificity, these values are higher than those reported in the present study. However, JoinMap 4.0 showed more common markers and greater ordering resemblance to the physical map than GACD, indicating more similarity to the actual scenario portrayed by the physical map.

The availability of genetic maps, linkage maps, and molecular information for *P. guajava* improves breeding programs and helps identify QTL (quantitative trait loci) linked to markers. Some QTL related to fruit traits (fruit morphology, fruit weight, and quality) and plant height were already reported and can be used in marker-assisted selection when linked to molecular markers (Ritter et al., 2010b). Dense genetic maps are crucial for that purpose (Nimisha et al., 2013). However, even though genetic maps are extensively applied in the scientific literature and are essential for QTL analysis and marker-assisted selection, they are still prone to error throughout several steps of the process and are expensive. Therefore, linkage analysis compared to physical maps based on sequence alignment is necessary to minimize algorithm errors or inconsistencies (especially when grouping and ordering markers).

There is still limited information on QTL mapping turning into finished plant breeding products through marker-assisted selection due to issues such as small to moderate-sized mapping populations, screening with a relatively small number of markers, and the difficulty to experimentally validate the roles of genes in the quantitative traits of many species (Evans et al., 2021). We strongly believe that another critical factor to this low conversion rate (QTL mapping into cultivar development) rests on the lack of homology between linkage and physical maps, as reported in our study.

Therefore, the physical map proposed here can be used as a reference for genetic mapping and QTL analysis in *P. guajava* due to its easy application, low cost, and, especially, the potential to generate information at the nucleotide level in the genome. Furthermore, marker association to the actual position in the genome and the anchoring of new markers can help identify candidate genes in QTLs for marker-assisted selection.

CONCLUSIONS

The genetic linkage maps spanned 1405.2 cM and 1392.7 cM, with average marker distances of 7.7 cM and 8.8 cM for JoinMap and GACD, respectively, and no agreement between the estimated distances and SNP ordering in the linkage groups. Both maps showed linkage groups with segments from several chromosomes compared to the physical map, indicating limitations. GACD showed greater limitation than JoinMap 4.0 when ordering markers according to their parental origin.

The physical map generated with BLAST consisted of 694 SNPs spanning 434.88 Mb, with an average marker distance of 0.62 Mb. The physical map reported here is useful for guava breeding, serving as a reference for future mapping studies and QTL estimates.

ACKNOWLEDGMENTS

The authors thank the National Council for Scientific and Technological Development (CNPq) of Brazil for the financial support for this study. Thanks to Dr. Dario Grattapaglia for providing the guava population genotyping data.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol.* 215: 403-410.

Blair MW, Cortés AJ, Farmer AD, Huang W, et al. (2018). Uneven recombination rate and linkage disequilibrium across a reference SNP map for common bean (*Phaseolus vulgaris* L.). *PLoS One.* 1: e0189597.

Bueno MRP (2009). O Projeto Genoma Humano. *Revista Bioética.* 5: 1-10.

Chagné D, Ryan J, Saeed M, Van Stijn T, et al. (2019). A high density linkage map and quantitative trait loci for tree growth for New Zealand mānuka (*Leptospermum scoparium*). *N. Z. J. Crop. Hortic. Sc.* 47: 261-272.

DeWan AT, Parrado AR, Matise TC and Leal SM (2002). The map problem: a comparison of genetic and sequence-based physical maps. *Am. J. Hum. Genet.* 70: 101-107.

Evans KS, van Wijk MH, McGrath PT, Andersen EC, et al. (2021). From QTL to gene: facilitates discoveries of the genetic mechanisms underlying natural variation. *Trends Genet.* 37: 933-947.

Feng C, Feng C, Lin X, Liu S, et al. (2020). A chromosome-level genome assembly provides insights into ascorbic acid accumulation and fruit softening in guava (*Psidium guajava*). *Plant Biotechnol. J.* 19: 717-730.

Gao W, Qu J, Zhang J, Sonnenberg A, et al. (2018). A genetic linkage map of *Pleurotus tuoliensis* integrated with physical mapping of the de novo sequenced genome and the mating type loci. *BMC genomics*. 19: 18.

Grattapaglia D, Mamani EMC, Silva-Junior OB and Faria DA (2015). A novel genome-wide microsatellite resource for species of Eucalyptus with linkage-to-physical correspondence on the reference genome sequence. *Mol. Ecol. Resour.* 15: 437-448.

Hohmann U, Endo TR, Gill KS and Gill BS (1994). Comparison of genetic and physical maps of group 7 chromosomes from *Triticum aestivum* L. *Mol. Gen. Genet.* 245: 644-653.

Inglis PW, Pappas MCR, Resende LV and Grattapaglia D (2018). Fast and inexpensive protocols for consistent extraction of high-quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS One.* 13: e0206085.

Jiangtao C, Yingzhen K, Qian W, Yuhe S, et al. (2015). MapGene2Chrom, a tool to draw gene physical map based on Perl and SVG languages. *Hereditas*. 37: 91-97.

Kosambi DD (1943). The estimation of map distances from recombination values. *Ann. Eugen.* 12: 172-175.

Lander ES, Linton LM, Birren B, Nusbaum C, et al. (2001). Initial sequencing and analysis of the human genome. *Nature.* 409: 860-921.

Lee YR, Yoon JB and Lee J (2016). A SNP-based genetic linkage map of *Capsicum baccatum* and its comparison to the *Capsicum annuum* reference physical map. *Mol. Breed.* 36: 61.

Lepitre V, Nansot G, Grangeon R, Pomies V, et al. (2010). The microsatellite (SSR)/AFLP reference linkage map of guava. *Acta Hortic.* 849: 183-192.

Masud Parvez GM, Uzzaman S, Miah Akanda K and Mehjabin S (2018). A short review on a Nutritional Fruit: Guava. *Toxicol. Res.* 1: 1-8.

Myburg A, Grattapaglia, D, Tuskan G, et al. (2014). The genome of *Eucalyptus grandis*. *Nature.* 510: 352-362.

Nimisha S, Kherwar D, Ajay KM, Singh B, et al. (2013). Molecular breeding to improve guava (*Psidium guajava* L.): Current status and future prospective. *Sci. Hortic.* 164: 578-588.

Padmakar B, Kanupriya C, Latha PM, Prashant KS, et al. (2015). Development of SRAP and SSR marker-based genetic linkage maps of guava (*Psidium guajava* L.). *Sci. Hortic.* 192: 158-165.

Paiva JAP, Prat E, Vautrin S, Santos MD, et al. (2011). Advancing Eucalyptus genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries. *BMC Genomics*. 12: 137-150.

Risterucci AM, Duval MF, Rohde W and Billotte N (2005). Isolation and characterization of microsatellite loci from Psidium guajava L. *Mol. Ecol. Notes*. 5: 745-748.

Ritter E, Herran A, Valdés-Infante J, Rodríguez-Medina NN, et al. (2010a). Comparative linkage mapping in three guava mapping populations and construction of an integrated reference map in guava. *Acta Hortic.* 849: 175-182.

Ritter E, Rodríguez-Medina NN, Velásquez B, Rivero D, et al. (2010b). QTL (quantitative trait loci) analysis in guava. *Acta Hortic.* 849: 193-202.

Rodríguez N, Valdés-Infante J, Becker D, Velázquez B, et al. (2007). Characterization of guava accessions by ssr markers, extension of the molecular linkage map, and mapping of qtls for vegetative and reproductive characters. *Acta Hortic*. 735: 201-216.

Silva-Junior OB, Faria DA and Grattapaglia D (2015). A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytol.* 206: 1527-1540.

Sumathi M, Bachpai VKW, Mayavel A, Dasgupta MG, et al. (2018). Genetic linkage map and QTL identification for adventitious rooting traits in red gum eucalypts. *3 Biotech.* 8: 242-252.

Van Ooijen JW (2006). JoinMap®4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Kyazma BV, Wageningen.

Van Ooijen JW and Jansen J (2013). Genetic mapping in experimental populations. Cambridge University Press, Cambridge.

Zhang L, Meng L, Wu W and Wang J (2015). GACD: Integrated software for genetic analysis in clonal fl and double cross populations. *J. Hered*. 106: 741-744.

Wu R, Ma C, Painter I and Zeng Z (2002). Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. *Theor. Popul. Biol.* 61: 349-363.

