Quantitative Trait Loci (QTL) Analysis of Fruit and Agronomic Traits of Tropical Pumpkin (Cucurbita moschata) in an Organic Production System

Angel R. Del Valle Echevarria 1, Alexandra Campbell 1, Theodore J. K. Radovich 1, Tia Silvasy 1, Sarah Moore 2 and Michael B. Kantar 1,*

1 Department of Tropical Plants and Soil Sciences, University of Hawai‘i–Mānoa, 3190 Maile Way, Honolulu, HI 96822, USA; angeldve@hawaii.edu (A.R.D.V.E.); colferab@hawaii.edu (A.C.); theodore@hawaii.edu (T.J.K.R.); tiameer@gmail.com (T.S.)
2 Department of Agronomy and Plant Genetics, University of Minnesota, 1991 Upper Buford Circle, St. Paul, MN 55108, USA; moor1579@umn.edu
* Correspondence: mbkantar@hawaii.edu

Received: 5 January 2020; Accepted: 24 February 2020; Published: 2 March 2020

Abstract: Interest in the development of organically grown vegetable crops has risen over the past decades due to consumer preferences. However, most crops that have desirable consumer traits have been bred in conventional growing conditions, and their transfer to an organic setting is challenging. Here, the organically grown Hawaiian pumpkin (Cucurbita moschata) accession ‘Shima’ was crossed with the conventionally grown Puerto Rican variety ‘Taina Dorada’ to develop a backcross (BC\textsubscript{1}) population, where ‘Shima’ was the recurrent parent. A total of 202 BC\textsubscript{1} (‘Shima’ X F\textsubscript{1}) progenies were planted in a certified organic field, and twelve traits were evaluated. We used genotype-by-sequencing (GBS) to identify the Quantitative Trait Loci (QTL) associated with insect tolerance along with commercially desirable traits. A total of 1582 single nucleotide polymorphisms (SNPs) were identified, from which 711 SNPs were used to develop a genetic map and perform QTL mapping. Reads associated with significant QTLs were aligned to the publicly available Cucurbita moschata genome and identified several markers linked to genes that have been previously reported to be associated with that trait in other crop systems, such as melon (Cucumis melo L.). This research provides a resource for marker-assisted selection (MAS) efforts in Cucurbita moschata, as well as serving as a model study to improve cultivars that are transitioning from a conventional to an organic setting.

Keywords: insect tolerance; color; marker assisted selection; qtl mapping

1. Introduction

Conventional breeding efforts have focused on selecting in high levels of inorganic fertilizer and crop protection inputs. Consequently, crops that are bred in conventional farming practices often struggle when they are transferred to low-input organic systems [1]. Organic farming looks for resilient crops that have adapted to agroecosystems containing different levels of inputs, diverse soil microorganisms, increased weed pressure, and insect pressure [2]. Plant breeding efforts in organic farming conditions have become increasingly popular due to consumer preference and the urgency of producing food in somewhat unstable environments, while still maximizing its yield [3].

The use of molecular markers has become a viable way to identify genetic loci associated with desired phenotypes in parallel with successful establishment in an organic setting [4,5]. However, most of these approaches focus on major crops such as wheat [6], sweet corn [4], tomato [7], and barley [8],
and few studies are focused on smaller crops such as pumpkin (Cucurbita moschata: $2n = 2x = 40$). Hawai‘i presents a unique opportunity to focus on molecular characterization for plant breeding on small crops due to its fertile soils [9], year-round growing environment [10], and consumer demand for organic products [11].

Tropical pumpkin (Cucurbita moschata) is an important crop worldwide due to its nutritional value and wide adaptability [12]. Unfortunately, pumpkin is susceptible to insect damage from the Lepidoptera species, mainly Diaphania nitidalis (pickleworm) [13]. Pickleworm is major cucurbit pest [14], which is very common worldwide and in Hawai‘i [15]. Pickleworm burrows into the blossom, as well as the immature fruit, negatively impacting the overall yield, fruit quality, and perishability of the mature fruit [14]. Biocontrol measures against pickleworm provides mixed results and often is not reliable [16].

‘Shima’ is a locally collected accession observed to have some pickleworm tolerance and has a slightly ribbed, flattened shape with yellowish flesh and weighing between 1–3 kg, traits desired by southeast Asian consumers but lacking broader market acceptance. ‘Taina Dorada’ is a cultivar developed in Puerto Rico [17] that has a smooth, spheroid-to-oblative spheroid shape, and high-quality flesh that is darker orange than ‘Shima’, with fruit weighing about twice that of ‘Shima’. The darker orange color of ‘Taina Dorada’ is more broadly preferred in some markets than the yellowish flesh of ‘Shima’, while the putative insect tolerance of ‘Shima’ has obvious benefits for organic systems.

The purpose of this study was to determine genetic associations on pumpkin fruit marketability and insect tolerance in a certified organic production setting. These two parental lines that show insect tolerance (‘Shima’) and consumer appeal (‘Taina Dorada’) were used to develop a mapping population. A backcross population with ‘Shima’ as the recurrent parent was genotyped using genotyping-by-sequencing (GBS) to identify Quantitative Trait Loci (QTL) associated with insect tolerance and fruit marketability. Our study has two main contributions: (1) seed that can be used to develop new lines that are insect tolerant and appealing to the consumer and (2) give genetic information that can be used to study other similar traits in organic environments for vegetables transitioning from conventional to organic farming.

2. Materials and Methods

2.1. Development of the Mapping Population

‘Shima’ was crossed to ‘Taina Dorada’ as the seed parent to develop the F_1 population. The F_1 was then crossed as the pollen parent to ‘Shima’. The BC_1 population was grown at the University of Hawai‘i Waimanalo Agricultural Research Station (see Field Conditions).

2.2. Field Conditions

The University of Hawai‘i Waimanalo Agricultural Research Station (UH-WARS) is located at the island of Oahu (21.3355° N, 157.7148° W) in the State of Hawai‘i. This tropical environment has varying temperatures from 20–29 degrees Celsius (68–84 degrees Fahrenheit), an average annual rainfall of 140 cm (55 in), and a mollisols soil type. The BC_1 population was planted on 9 September 2016 (96 individual) and 5 January 2017 (106 individuals). The organic plots were 15 × 30.5 m (50 × 100 ft), and standard organic practices were implemented. Tropical pumpkin has about a 150 day growing period, and flowering time is approximately eight weeks. Seedlings were planted in the UH-WARS greenhouse and transplanted to the field at two weeks of age. We applied granular organic fertilizer (Sustane; Sustane Corp., Cannon Falls, MN at a rate of 150 lbs/acre) two weeks before planting and incorporated it into the top 6” of soil. An organic pesticide rotation program was implemented starting at the 5th leaf growth stage until flowering was complete, applied every 10 days. Fruit fly traps were hung in empty modified 2-L plastic bottles at time of flowering and fruit set. Weed control was performed by hand hoeing.
2.3. Evaluation of Traits

A total of twelve traits were evaluated in the BC$_1$ population. These traits include: fresh pumpkin weight, *Diaphania nitidalis* (pickleworm) damage, pumpkin height, cavity volume, thickness of the flesh outside and inside the rib, height of the pumpkin cavity inside and outside rib, cavity width, overall flesh color intensity, flesh red color intensity, and flesh yellow color intensity. Fresh pumpkin weight was measured in pounds (lbs) using a weight scale. Pickleworm damage was visually determined by using the following rubric: 0 = no damage, 1 = some damage (~one worm scar), 2 = two scars, 3 = three scars, and 4 = four or more scars. Flesh thickness, cavity height, and cavity width measurements were done with a ruler in cm. Color intensity measurements were used with a Nix color sensor (Nix Sensor Ltd., Hamilton, Ontario,) using black (overall intensity), red (red intensity), and yellow (yellow intensity) cards to calibrate the colorimeter to then measure the pumpkin flesh.

2.4. DNA Extraction and Genotype-by-Sequencing (GBS)

Genomic DNA was extracted from leaf tissue BC$_1$ individuals using the E.N.Z.A. SP Plant DNA Kit (Omega BioTek, Norcross, GA, USA) following the manufacturer’s recommendations. DNA was quantified and assessed for quality by using Nanodrop 1000 Spectrophotometer V3.8 (Thermo Scientific, Wilmington, DE, USA). DNA samples were sent to Data2Bio, LLC (Ames, IA, USA) for tGBS® Genotype by Sequencing technology (2-bp selection); they provided a full report of sequencing, single nucleotide polymorphisms (SNP) calling, and linkage map creation using R/qtl [18,19]. Briefly, Data2Bio generated consensus sequences that could be used to call SNPs without a reference by using quality trimmed reads from all samples, which were combined and normalized to a maximum of 50× coverage. Briefly, quality trimmed reads from all samples were combined and normalized to a maximum of 50× coverage. The sequencing errors in the reads were then corrected using Fiona [20]. Coverage-normalized and error-corrected reads were then condensed using CD-HIT-454 with ≥96% identity to form consensus clusters. Clusters with <10 component reads and <50 bp in length were discarded. A total of 356,440 consensus sequences were obtained with a total sequence length of 35.5 Mb. SNPs were filtered using stringent criteria, where the minimum call rate was ≥50%, the allele number was 2, the number of genotypes was ≥2, the minor Allele Frequency (MAF) was >10%, and the heterozygosity rate was between 25 and 75%. Raw reads are available within the NCBI Short read achieve with project number SRX4484561.

2.5. Linkage Map Development and Quantitative Trait Loci (QTL) Mapping

In order to create the linkage map, the following parameters were used in R/qtl software, a recombination frequency of 0.35, a LOD cut off of 8, and a minimum number of SNPs per linkage group of 10, with the Haldane mapping function. While there are 20 chromosomes, attempts to combine linkage groups to match the number of chromosomes were unsuccessful as there were not enough linked markers to achieve this. R/qtl was used in R software version 3.5.0 [18] following Broman et al. [19] suggestions for QTL identification. All traits were evaluated for the identification of QTLs based on their phenotypic distribution (Figures S1–S3) using the “scanone” function, in which 10,000 permutations were performed using the “n.perm” argument to determine the significance threshold at alpha = 0.05. Traits that showed a normal distribution such as fresh pumpkin weight, pumpkin height, thickness of the flesh outside and inside the rib, and height of the pumpkin cavity inside and outside rib were analyzed using the “hk”, Haley–Knott regression method (Figure S1). Traits that showed bimodal distribution such as cavity width, overall flesh color intensity, flesh red color intensity, and flesh yellow color intensity were analyzed using the “2part” method (Figure S2). Finally, traits that showed a nonparametric distribution, such as *Diaphania nitidalis* (pickleworm) damage and cavity volume, were analyzed using the “np” method (Figure S3). Traits that showed peaks that were greater than the calculated threshold were determined as significant, and markers were identified using the “summary” function that had the permutation (perm) and alpha (0.05) arguments. Coordinates in
the genetic map are provided, and by using the “find.marker” function, the significant read (or SNP) was identified.

2.6. Genomic Analysis

Individual GBS reads that contained a SNP that showed a statistically significant association with any given trait in the QTL analysis were aligned against the pumpkin genome [21] that is publicly available through the Cucurbit Genomics Database [22] using the BLAST tool [23]. Coordinates and annotations are summarized in Table S1. GBS marker reads that seem to have a positive hit on a gene model were compared with the melon (*Cucumis melo* L.) genome [24] to identify possible gene functions and find associations with the evaluated traits of this study (Table S2).

3. Results and Discussion

3.1. Genotype-by-Sequencing

A total of 560.8 million reads were processed from the BC$_1$ plants after sequencing with an average of 2.8 million reads per sample. After quality filter checks, 483.8 million reads at 106 bp in length were obtained with an average of 2.4 million reads per sample. A total of 356,440 consensus sequences were generated that could be used in lieu of a reference genome for alignment and SNP calling. Subsequently, all SNPs identified were filtered to explore different missing data rates used for genotyping across the samples. The resulting numbers of SNPs remaining, missing data points, and detected polymorphism rate for each minimum calling rate (MCR) level are displayed in Table 1.

**Table 1.** Single nucleotide polymorphisms (SNP) genotyping summary. MCR50 was selected as the appropriate cut-off for this study. At this cut-off, the % polymorphism was ~0.18%.

<table>
<thead>
<tr>
<th>Minimum Calling Rate (MCR)</th>
<th>No. SNPs</th>
<th>Missing Data Points</th>
<th>% Polymorphisms *</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCR50</td>
<td>1582</td>
<td>82,434/314,818 = 26.2%</td>
<td>1582/859,838 = 0.18%</td>
</tr>
<tr>
<td>MCR60</td>
<td>1101</td>
<td>39,111/219,099 = 17.9%</td>
<td>1101/699,932 = 0.16%</td>
</tr>
<tr>
<td>MCR70</td>
<td>805</td>
<td>18,318/160,195 = 11.4%</td>
<td>805/578,022 = 0.14%</td>
</tr>
<tr>
<td>MCR80</td>
<td>617</td>
<td>8861/122,783 = 7.2%</td>
<td>617/460,450 = 0.13%</td>
</tr>
<tr>
<td>MCR90</td>
<td>415</td>
<td>3069/82,585 = 3.7%</td>
<td>415/346,589 = 0.12%</td>
</tr>
</tbody>
</table>

3.2. Linkage Map

Subsequently, a subset of 1582 SNPs that had less than 50% missing data across the BC$_1$ population was used to construct a genetic map from the 202 individuals in the population. Of the 1582 MCR50 SNPs, 711 (i.e., 44.9%) were successfully mapped. The resulting genetic map consisted of 24 linkage groups with an overall genetic length of 2665 cM and an average spacing of 3.7 cM (Figure 1). The mapping was unable to achieve a chromosome level resolution as in recent studies [21] due to an insufficient number of markers, but the map did provide close to a chromosome-level number of linkage groups that provide utility in tropical pumpkin breeding.
3.3. QTL Mapping and Genomic Analysis.

The 202 BC₁ progeny were evaluated for 12 agronomically important traits relevant to marketability and organic production (Table 2). We observed variability among the traits, so we analyzed their distribution to determine the appropriate model for QTL analysis. Four traits showed statistically significant QTLs across the pumpkin genome (Figure 2). Among these four traits, eight markers were associated with at least one trait, and one marker, Pumpkin_164349 in linkage group 9, was associated with three of the four traits (Table S1). We used the publicly available pumpkin genome to evaluate the reads that were significantly associated with the evaluated traits (Table S1; [22–24]). Two markers, Pumpkin_89537 and Pumpkin_216447, showed multiple hits in the pumpkin genome, very likely associated with repetitive regions in the genome, and were disregarded for further analysis. Six of the eight markers were homologous to a region in the pumpkin genome that had a gene model. We took the sequence of these gene models and compared them to the Melon genome [25] by sequence analysis (Tables S1 and S2; [23,24]). We decided to analyze our data against the melon genome, because there have been several QTL and genomic studies on melon fruit quality that are similar to our study [26–28].

Markers associated with fruit flesh color intensity Pumpkin_282381, Pumpkin_93626, Pumpkin_164349, Pumpkin_46427, Pumpkin_210956, and Pumpkin_221779 aligned with pumpkin gene models whose orthologs in melon have been measured to be differentially regulated in green versus orange fruit color [26] and fruit flesh tissue [27]. In parallel, Pumpkin_164349, which our study showed to be associated with pickleworm tolerance, aligns with CmoCh15G001520, a clathrin heavy chain protein encoding gene orthologous to MELO3C008306 that has been observed to be upregulated in Fusarium infections [28]. These markers may serve for future molecular breeding efforts to develop cultivars well adapted to organic production systems with broad consumer acceptance.
Figure 2. Significant QTLs identified in the BC1 pumpkin population. (A) *Diaphania nitidalis* (pickleworm) damage visible on the pumpkin surface showed one peak in linkage groups: 1, 9, and 15. (B) Cavity width showed one peak in linkage group 1. (C) Flesh color intensity showed one peak in linkage groups 1, 9, 12, and 17. (D) Red color intensity in the fruit flesh showed one peak in linkage groups 1, 9, 12, 15, and 17. Each of the analyses implemented 10,000 permutations to establish a threshold at alpha = 0.05.

Table 2. Traits measured in the BC1 population.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean ±/− Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fresh pumpkin weight</td>
<td>4.8 ±/− 0.18 lbs</td>
</tr>
<tr>
<td><em>Diaphania nitidalis</em> (pickleworm) damage</td>
<td>0.9 ±/− 0.08 z</td>
</tr>
<tr>
<td>Cavity volume</td>
<td>392.4 ±/− 17.09 mL</td>
</tr>
<tr>
<td>Thickness of the flesh outside the rib</td>
<td>3.3 ±/− 0.09 cm</td>
</tr>
<tr>
<td>Thickness of the flesh inside the rib</td>
<td>3.2 ±/− 0.09 cm</td>
</tr>
<tr>
<td>Height of the pumpkin cavity outside rib</td>
<td>18.5 ±/− 0.25 cm</td>
</tr>
<tr>
<td>Height of the pumpkin cavity inside rib</td>
<td>17.6 ±/− 0.22 cm</td>
</tr>
</tbody>
</table>
Table 2. Cont.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean ±/− Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavity width</td>
<td>11.7 ±/− 0.51 cm</td>
</tr>
<tr>
<td>Overall flesh color intensity</td>
<td>63.3 ±/− 1.19 y</td>
</tr>
<tr>
<td>Flesh red color intensity</td>
<td>22.0 ±/− 1.20 x</td>
</tr>
<tr>
<td>Flesh yellow color intensity</td>
<td>46.1 ±/− 2.34 w</td>
</tr>
</tbody>
</table>

* Visual evaluation: 0 = no damage, 1 = some damage (~one worm scar), 2 = two worm scars, 3 = three worm scars, and 4 = four or more worm scars. y Colorimeter calibrated with a black background: 0–50 is a darker color and 51–100 is a lighter color intensity. x Colorimeter calibrated with a cherry red piece of paper, where negative values were on the green side of the spectrum and positive values were on the red side of the spectrum. w Colorimeter calibrated with a yellow piece of paper, where negative values were on the blue side of the spectrum and positive values were on the yellow side of the spectrum.

4. Conclusions

Although we focused on molecular pumpkin breeding from conventional to organic farming in Hawai‘i, this strategy can be applied to any crop whose main production occurs in a conventional setting and consumers demand an organically grown alternative. Indeed, we found that new technologies were as successful in this understudied system as in others [29,30]. Furthermore, our genetic and genomic information can be used to improve the genome assembly of pumpkin. Initially, we compared our 1582 marker reads against the genome and identified that 325 (20%) marker reads mapped to chromosome 0 (data not shown). We identified large effect QTL for agronomic and quality traits that are important for the Hawaiian pumpkin market. The populations and markers developed here are currently being used in the further development of tropical pumpkin. Further, studies using segregating populations, GBS, and higher sequence depth could produce high-quality genomes that could serve for future genetic and genomic studies for crop improvement.

Supplementary Materials: The following are available online at http://www.mdpi.com/2311-7524/6/1/14/s1, Figure S1: Phenotypic distribution of traits evaluated by the Harley-Knott (hk) regression method in R/qtl, Figure S2: Phenotypic distribution of traits evaluated by 2-part, also known as bimodal, method in R/qtl, Figure S3: Phenotypic distribution of traits evaluated by non-parametric method in R/qtl. Table S1: Sequence analysis between significant reads in this study and the pumpkin genome, Table S2: Table S2. Genomic comparison between *C. moschata* gene models and *Cucumis melo* genome.

Author Contributions: Conceptualization, M.B.K. and T.J.K.R.; methodology, S.M., A.C., T.S.; formal analysis, A.C., A.R.D.V.; investigation, data curation, M.B.K., A.R.D.V.; writing—original draft preparation, A.C., A.R.D.V.; writing—review and editing, M.B.K., T.J.K.R., S.M., T.S.; project administration, S.M., T.S.; funding acquisition, M.B.K., T.J.K.R. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by USDA-ARS cooperative agreement 58-5320-4-020 the Control of Pests and Diseases and Adding Value to Specialty Crops.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References


13. Elsey, K.D. Resistance mechanisms in *Cucurbita moschata* to pickleworm and melonworm (Lepidoptera: Pyralidae). *J. Econ. Entomol.* 1985, 78, 1048–1051. [CrossRef]


