



## OPEN QTL associated with resistance to *Stagonosporopsis citrulli* in *Citrullus amarus*

Lincoln Adams & Cecilia McGregor<sup>✉</sup>

Gummy stem blight (GSB) is a fungal disease affecting cucurbit crops, including watermelon (*Citrullus lanatus*), leading to significant yield losses. The disease is caused by three *Stagonosporopsis* species, of which *Stagonosporopsis citrulli* is the most common in the southeastern United States. Currently no gummy stem blight-resistant watermelon cultivars are available to growers. In this study, QTL-seq in an interspecific population developed from Sugar Baby × PI 189225 (*Citrullus amarus*) identified QTL on chromosomes 2, 5, 9 and 11. A novel QTL on chromosome 5 (*Qgsb5.2*) associated with resistance to *S. citrulli* (PVE = 13.3%) was confirmed by genetic mapping. KASP marker assays were developed for selection of *Qgsb5.2* to allow breeders to track the allele contributing resistance to GSB, reducing the need for laborious phenotyping. Pyramiding different GSB resistance QTL could be a useful strategy to develop GSB resistant watermelon cultivars.

Gummy stem blight (GSB) is a fungal disease which affects cucurbit crops globally, leading to decreases in yield and revenue loss for growers<sup>1,2</sup>. Watermelon (*C. lanatus*) is one of the most severely affected crops with average yield losses of up to 43% in unsprayed plots<sup>3</sup>. Until recently it was believed that GSB was caused by a single pathogen, *Didymella bryoniae*. However, Stewart et al.<sup>2</sup> determined that GSB is caused by three distinct species of *Stagonosporopsis*, *S. cucurbitacearum*, *S. caricae* and *S. citrulli*<sup>2</sup>. Currently, GSB is controlled using costly fungicides that are damaging to the environment highlighting the need for resistant cultivars<sup>4,5</sup>.

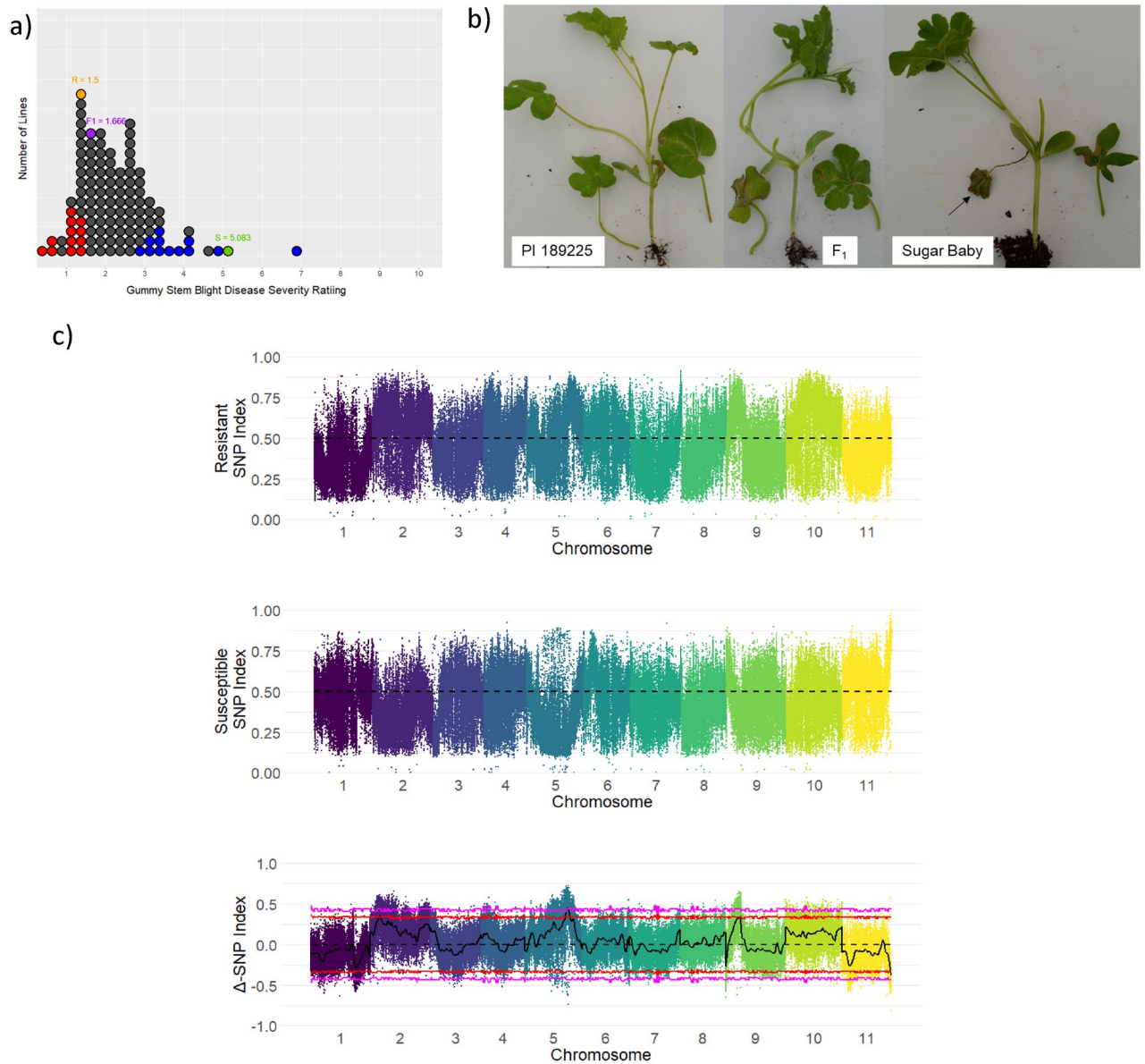
Initial efforts to breed watermelon with GSB resistance were conducted working under the hypothesis that the resistance in PI 189225 (*C. amarus*) was controlled by a single gene, designated *db*<sup>6</sup>. Norton used two resistant accessions, PI 189225 and PI 271778 (*C. lanatus*), to develop four cultivars ('AU-Producer', 'AU-Jubilant', 'AU-Sweet Scarlet', and 'AU-Golden Producer') with reported GSB resistance<sup>7-9</sup>. Unfortunately, these varieties did not exhibit the same level of GSB resistance once they were grown by producers on a larger scale<sup>10</sup>. Additional sources of GSB-resistant watermelon germplasm have been identified by other researchers in recent years<sup>11-13</sup>; however, to date, none of these discoveries have led to the release of GSB-resistant commercial watermelon cultivars. Recently, Gusmini et al.<sup>14</sup> demonstrated that GSB resistance in four crosses between elite watermelon cultivars and resistant PIs, including PI 198225, is quantitatively inherited<sup>14</sup>.

In three previous studies, researchers identified GSB resistance QTL in watermelon. Ren et al.<sup>15</sup> used a population derived from K3 (*C. lanatus*) × PI 189225 and an isolate of *S. cucurbitacearum* to map a GSB resistance QTL on chromosome 8<sup>15</sup>. Lee et al.<sup>16</sup> used a population derived from 920533 (*C. lanatus*) × PI 189225 and "*D. bryoniae* 'KACC 40937 isolate'" to map GSB resistance QTL, two on chromosome 8 and one on chromosome 6<sup>16</sup>. The QTL mapped by the two studies on chromosome 8 are distinct from one another. The species of *Stagonosporopsis* used by Lee et al.<sup>16</sup> is unknown. Gimode et al.<sup>17</sup> used a population derived from Crimson Sweet (*C. lanatus*) × PI 482276 (*C. amarus*) and *S. citrulli* isolate 12178A to map GSB resistance QTL on chromosomes 5 and 7<sup>17</sup>. Differences in the locations of these QTL, especially those derived from the same resistant parent (PI 189225), are likely due to either differences in resistance to the various species or isolates of *Stagonosporopsis* used for phenotyping or differences in phenotyping methodologies. Lee et al.<sup>16</sup> evaluated lesion severity on the stems separately from leaf lesions, while Ren et al.<sup>15</sup> scored only leaf lesions and Gimode et al.<sup>17</sup> scored the entire seedling<sup>15-17</sup>.

QTL-seq, a modification of bulked segregant analysis, was proposed by Takagi et al.<sup>18</sup> as a quick and relatively cheap method to identify QTL associated with a particular trait<sup>18,19</sup>. This technique has been widely used to identify QTL for various traits, including blast resistance in rice, late spot resistance in peanut, and Fusarium wilt and GSB resistance in watermelon<sup>15-18,20-22</sup>.

The aim of the current study was to use QTL-seq to identify loci associated with resistance to *S. citrulli*, the most common GSB-causing *Stagonosporopsis* species in the southeastern US, in a population derived from a

Institute of Plant Breeding, Genetics, and Genomics, University of Georgia, Athens, GA, USA. ✉email: cmcgre1@uga.edu



**Figure 1.** (a) Stacked dot plots showing the distribution of the average disease severity of the  $F_{2:3}$  families ( $n = 111$ ) of a cross between Sugar Baby (S) and PI 189225 (R). Red and blue dots represent the families selected to generate the resistant and susceptible DNA bulks, respectively. The orange, purple, and green dots represent the average phenotypic values for PI 189225, the  $F_1$ , and Sugar Baby, respectively. (b) Gummy stem blight disease severity in PI 189225 (resistant), the  $F_1$ , and Sugar Baby (susceptible) after infection with *S. citrulli*. The black arrow indicates a completely dead leaf in the susceptible parent. (c) The SNP index for the gummy stem blight resistant and susceptible bulks and the  $\Delta$ -SNP index of each identified SNP. Individual colored points represent individual SNPs, with each color representing a different chromosome. The black line represents the tribune smoothed  $\Delta$ -SNP index value using a 1 Mb sliding window. The red line represents the 95% confidence interval and the pink line represent the 99% confidence interval. A positive  $\Delta$ -SNP index represents SNPs differing from the reference genome which are potentially contributing to the trait of interest, while regions with a negative  $\Delta$ -SNP index represent SNPs shared with the reference genome which are potentially contributing to the trait of interest. Regions where the black line exits the bounds of the confidence intervals represent potential QTL for gummy stem blight resistance.  $\Delta$ -SNP indices were calculated using QTLseqr<sup>48</sup>.

cross between Sugar Baby (SB, susceptible) and PI 189225 (resistant) and to develop KASP marker assays for selection of resistance loci in watermelon<sup>2,23</sup>.

## Results

**Phenotypic data.** The phenotypic distribution for the population was skewed (Shapiro–Wilk test  $P = < 0.0001$ ) towards lower disease symptom severity (Fig. 1a). PI 189225 (resistant parent), Sugar Baby (susceptible parent) and the  $F_1$  had disease severity scores of 1.5, 4.6 and 1.7, respectively (Fig. 1b). Broad sense herit-

Sample	Total reads	Mapped reads	Mapping ratio (%)	Properly paired (%)	Average coverage (x)
Susceptible bulk	148,037,545	145,179,598	98.07%	89.53%	56
Resistant bulk	149,620,933	147,148,116	98.35%	89.79%	57

**Table 1.** Mapping statistics for short reads generated by sequencing of the resistant and susceptible DNA bulks.

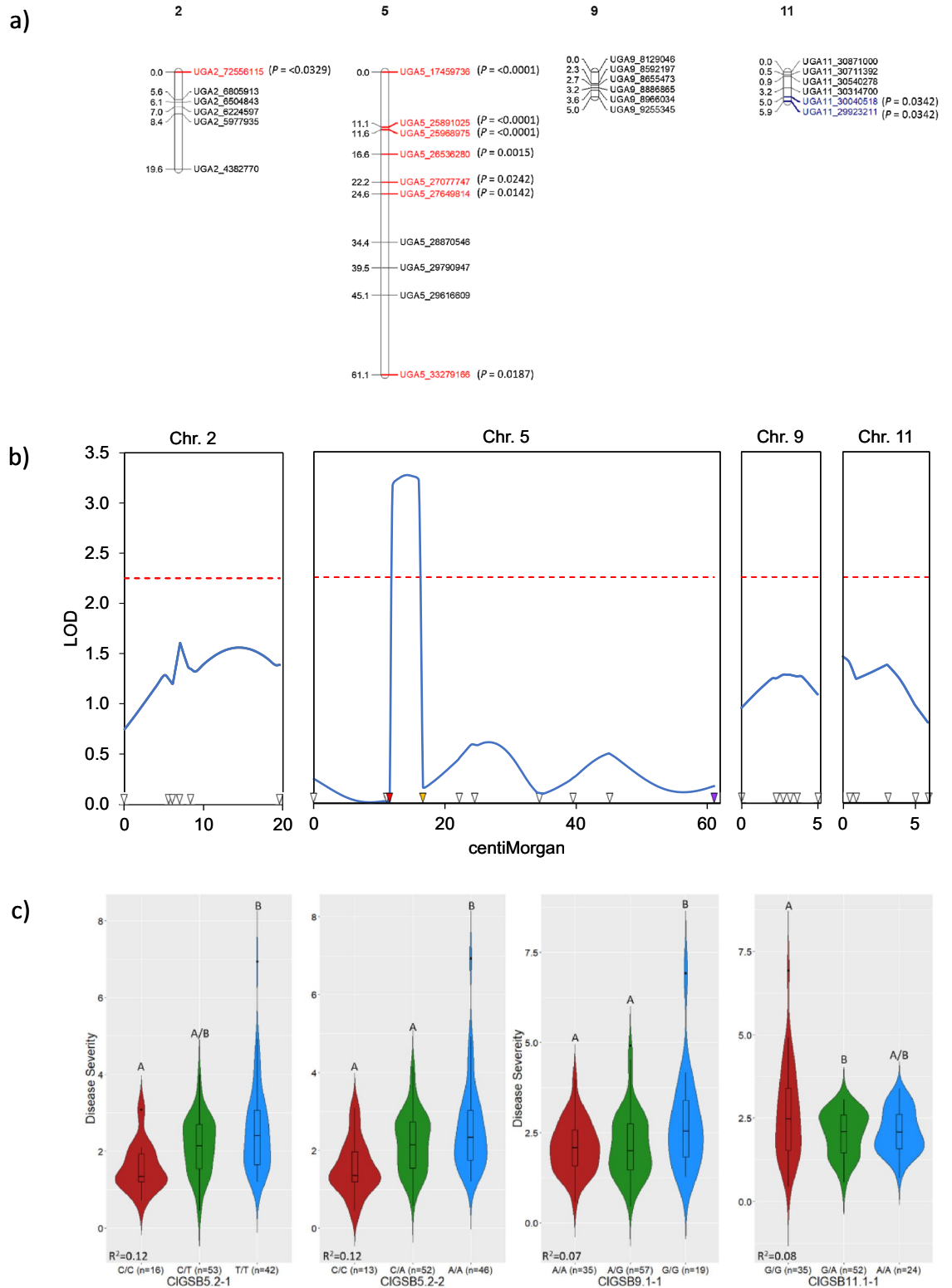
Chr	Start (bp)	End (bp)	Length (bp)	Total # of SNPs	Peak $\Delta$ -SNP index	Peak $\Delta$ -SNP index position
2	5,835,675	6,873,027	1,037,352	482	0.36	6,517,652
5	25,465,570	28,369,775	2,904,205	1677	0.45	26,916,783
9	8,570,733	8,949,900	379,167	174	0.34	8,843,908
11	30,375,412	30,879,210	503,798	147	-0.38	30,879,210

**Table 2.** Quantitative trait loci identified for gummy stem blight resistance by comparing the susceptible and resistant bulks of the Sugar Baby  $\times$  PI 189225  $F_{2:3}$  population. Locations are based on the *C. lanatus* 91703\_v2 genome sequence (Guo et al.<sup>24</sup>).

ability for disease severity was 0.27. The 20 most resistant and 20 most susceptible families from the first screen were re-screened (data not shown) and the 12 most resistant and susceptible plants across the two screens were used to create the resistant and susceptible bulks for QTL-seq. The average disease severity scores of the resistant and susceptible bulks over the two screens were 1.5 and 3.7, respectively.

**QTL-seq.** DNA sequences were obtained from each bulk consisting of 148,037,545 reads for the susceptible bulk and 149,620,933 reads for the resistant bulk. Alignment to the *C. lanatus* 91703\_v2 genome resulted in 145,179,598 mapped reads with a mapping ratio of 98.07%, with 89.53% of reads properly paired and an average coverage of 56 $\times$  for the susceptible bulk, and 147,148,116 mapped reads with a mapping ratio of 98.35% with 89.79% of reads properly paired and an average coverage of 57 $\times$  for the resistant bulk<sup>24</sup>. 6,186,663 SNPs were identified from the aligned reads (Table 1). Initial filtering with GATK reduced the number of SNPs to 5,806,149. The second round of filtering using QTLseqr filtered another 5,567,799 leaving 238,350 high quality SNPs for analysis.  $\Delta$ -SNP Index analysis using QTLseqr revealed four regions in which the tricube smoothed  $\Delta$ -SNP Index exceeded the 95% confidence threshold, one region each on chromosomes 2, 5, 9, and 11 (Fig. 1c, Table 2). The peak region on chromosome 2 exceeding the 95% confidence interval spanned 1,037,352 base pairs, from 5,835,675 to 6,873,027 bp, with a peak  $\Delta$ -SNP index of 0.36. The peak region on chromosome 5 exceeding the confidence interval spanned 2,904,205 base pairs, from 25,465,570 to 28,369,775 bp, with a peak  $\Delta$ -SNP Index of 0.45. The peak on chromosome 5 exceeded the 99% confidence interval. The peak region on chromosome 9 exceeding the 95% confidence interval spanned 379,167 base pairs, from 8,570,733 to 8,949,900 bp, with a peak  $\Delta$ -SNP Index of 0.34. The peak region on chromosome 11 exceeding the 95% confidence interval spanned 503,798 base pairs, from 30,375,412 to 30,879,210 bp, with a peak  $\Delta$ -SNP Index of -0.38.

**QTL mapping.** Twenty-eight KASP marker assays were developed spanning the chromosomal regions of interest. Genetic mapping resulted in four linkage groups spanning 19.61, 61.06, 5.01, and 5.90 cM, on chromosomes 2, 5, 9, and 11, respectively (Fig. 2a). Segregation distortion was observed for several markers on the map, predominantly in the direction of the susceptible SB parent (Fig. 2a). QTL mapping (Fig. 2b) resulted in a single significant peak on chromosome 5, between SNP UGA5\_25968975 (KASP assay ClGSB5.2-1) and UGA5\_26536280 (KASP assay ClGSB5.2-2) at 14.0 cM (LOD = 3.37; PVE = 13.3%), with left and right 1-LOD confidence intervals of 11.5 cM and 16.5 cM. Additive and dominance values for this QTL were 0.5414 and 0.1116, respectively. Significant preferential segregation of the susceptible alleles was observed for seven out of the ten markers mapped on chromosome five ( $\chi^2 P < 0.05$ ). To determine whether the QTL identified in the current study overlaps with *Qgsb5.1* (syn. *ClGSB5.1*) identified by Gimode et al.<sup>17</sup>, the marker most closely associated to *Qgsb5.1* (ClGSB5.1-1) was included in the map (Fig. 2b)<sup>17</sup>. ClGSB5.1-1 mapped 44.5 cM outside the confidence interval of the current QTL and it was therefore concluded that the QTL identified in the current study represents a novel QTL (*Qgsb5.2*). For the two KASP assays closest to *Qgsb5.2*, individuals homozygous for the allele from the susceptible parent (A/A or T/T) were significantly less resistant to GSB than individuals homozygous for the allele from the resistant parent (C/C) (Fig. 2c, Table 3). Assay ClGSB5.2-1 showed significant ( $P = 0.001$ ) association with disease resistance (C/C = 2.1; T/T = 2.9), with an  $R^2$  value of 12.1%. Assay ClGSB5.2-2 showed significant ( $P = 0.001$ ) association with disease resistance (C/C = 1.6; A/A = 2.6), with an  $R^2$  value of 12.3%. The haplotype representing both ClGSB5.2-1 and ClGSB5.2-2 showed a significant ( $P = 0.014$ ) association with disease resistance (C/C:C/C = 1.6; T/T:A/A = 2.6) for individuals homozygous at both loci for the resistant or susceptible alleles, with an  $R^2$  value of 12.2% (data not shown). The QTL identified by QTL-seq on chromosomes 2, 9, and 11 could not be confirmed by genetic mapping. However, single marker analysis found that individuals homozygous for the SB allele had significantly higher disease severity ( $P < 0.05$ ) than



heterozygotes or individuals homozygote for the PI 189225 alleles for all markers on chr. 9. On chr. 11, heterozygote individual had significantly lower disease severity than individuals homozygous for the PI 189225 alleles for four markers (UGA11\_30314700, UGA11\_30540278, UGA11\_30711392 and UGA11\_30871000). Individuals homozygous for the Sugar Baby allele, were not significantly different from heterozygotes or PI 189225 homozygotes. UGA9\_8886865 (KASP assay CIGSB9.1;  $P < 0.0106$ ;  $R^2 = 8.5\%$ ; AA = 2.1; AG = 2.2; GG = 2.9) and UGA11\_30871000 (KASP assay CIGSB11.1-1;  $P < 0.0203$ ;  $R^2 = 7.5\%$ ; GG = 2.7; GA = 2.18; AA = 2.1) were the

**Figure 2.** (a) Linkage maps of the regions on chromosomes 2, 5, 9, and 11 identified by QTLseq. These maps are comprised of 6, 10, 6 and 6 markers respectively with lengths of 19.6, 61.1, 5.0 and 5.9 cM. Markers indicated in red are preferentially segregating for the alleles contributed by SB, and markers indicated in blue are preferentially segregating for the alleles contributed by PI 189225. *P* values for Chi<sup>2</sup> goodness of fit test are given next to each marker. (b) QTL associated with Gummy stem blight resistance in the Sugar Baby × PI 189225 F<sub>2,3</sub> population (*n* = 111). The red dashed line indicates the LOD threshold (1000 permutations), while the triangles indicate the position of the KASP markers. The purple triangle represents marker CIGSB5.1-1, which was linked to *Qgsb5.1* in the Gimode et al. study<sup>17</sup>. The red triangle represents marker CIGSB5.2-1, and the orange triangle represents marker CIGSB5.2-2, both of which are linked to *Qgsb5.2*. (c) Violin plots showing the association of CIGSB5.2-2, CIGSB5.2-2, CIGSB9.1-1 and CIGSB11.1-1 genotypes with disease severity in the Sugar Baby × PI 189225 population (*n* = 111). Red plots represent individuals that are homozygous for the allele from the resistant parent (PI 189225) allele, while blue plots represent individuals homozygous for the susceptible parent (Sugar Baby) allele. Box and whisker plots show median and quartile ranges of each group. Different letters above the plots indicate significant differences (*P* < 0.05) based on a Tukey–Kramer test.

most highly significant marker on chrs 9 and 11, respectively (Fig. 2c). For the markers on chr. 11, heterozygous individuals had lower disease severity than the individual homozygous for the allele from the resistant parent.

**Syntenic analysis and candidate genes.** The 567,305 bp region of *Qgsb5.2* contained 38 predicted genes. Three of these genes, an Ethylene-responsive transcription factor 2 gene (*Cl97C05G096890*), a histidine kinase 5 gene (*Cl97C05G096980*) and an enhanced disease resistance 2-like protein gene (*Cl97C05G097030*), are probable fungal disease resistance genes based on their similarity to genes known to be related to fungal resistance in *Arabidopsis thaliana*<sup>24–27</sup>. No non-synonymous polymorphisms were detected in the exons of these genes within our sequencing data. The *Qgsb5.2* region was not found to be syntenic to any previously identified GSB resistance QTL in cucurbits<sup>15–17,28–30</sup>.

## Discussion

GSB causes significant yield losses for watermelon growers<sup>3</sup>. Efforts to control the disease using genetic resistance have recently been confounded by the discovery that the disease is caused by three *Stagonosporopsis* species<sup>2</sup>. A study by Gimode et al.<sup>31</sup> showed significant differences in disease severity caused by different *Stagonosporopsis* isolates in different watermelon genotypes, but no significant genotype × isolate interactions were observed<sup>31</sup>. Unraveling the host resistance response to different *Stagonosporopsis* species and isolates is likely key to developing cultivars with field resistance to this disease. Here we report a novel QTL from PI 189225 associated with resistance to *S. citrulli*, the most common species causing GSB in the southeastern US.

A continuous distribution was observed for disease symptom severity in the Sugar Baby × PI 189225 population, indicating quantitative control of GSB resistance in this population. Additionally, both resistant and susceptible families of the population exhibited transgressive segregation relative to the parents. This contributes additional evidence to the hypothesis that GSB resistance is quantitatively controlled<sup>14</sup>. Other recent studies have reached similar conclusions, in contrast to earlier reports which proposed resistance in PI 189225 was monogenic<sup>6,14,15,17,28–30,32–34</sup>.

QTL-seq identified QTL on chromosomes 2, 5, 9, and 11, and *Qgsb5.2* was confirmed by genetic mapping. Similar results were obtained by Ramos et al.<sup>35</sup> in a study mapping *Phytophthora capsici* resistance in squash<sup>35</sup>. Significant single marker associate was also found with markers on chromosomes 9 and 11, but not on chromosome 2. Possible explanations for the inability to confirm certain QTL, include the small population size and/or a small contribution to resistance by individual QTL<sup>36</sup>. The unconfirmed QTL could also have been a false positive.

The KASP assays (CIGSB5.2-1, CIGSB5.2-2, CIGSB9.1-1 and CIGSB11.1-1) were effective at explaining between 7 and 12% of the resistance observed in this population. While this shows that the markers can effectively track the loci, it also underlines the need for the identification of more resistance QTL, and to pyramid different QTL as a 12% increase in resistance would likely not lead to any appreciable reduction in fungicide use by farmers. The two markers most closely linked to *Qgsb5.2* (CIGSB5.2-1 and CIGSB5.2-2) exhibited severe segregation distortion (CIGSB5.2-1: *P* < 0.0001; CIGSB5.2-2: *P* = 0.0015) in the direction of the SB alleles. Segregation distortion was observed for all markers within 12 cM of *Qgsb5.2*. The significant segregation distortion is likely due to the interspecific nature of the cross which originated this population. Segregation distortion in *C. lanatus* × *C. amarus* crosses is a common phenomenon (Sandlin et al. 2012, Gimode et al.<sup>17</sup>) and complicates introgression of desirable alleles from *C. amarus* into elite watermelon.

Two previous studies using the same resistant source (PI 189225) identified two different QTL on chromosome 8 and one on chromosome 6<sup>15,16</sup>. A possible explanation for the different QTL identified in different studies is the use of different phenotyping methods or rating systems. The study by Lee et al.<sup>16</sup> rated stem lesions and leaf lesions separately and mapped both separately, while Ren et al.<sup>15</sup> rated plants based on the symptoms on four true leaves<sup>15,16</sup>. The current study rated plants based on the symptoms of the first two true leaves. Another possible explanation for the identification of different QTL is the use of different species of the GSB-causing fungi. Ren et al.<sup>15</sup> used an isolate of *S. cucurbitacearum*, while the current study was conducted using an isolate of *S. citrulli*<sup>15</sup>. The specific species of *Stagonosporopsis* used by Lee et al.<sup>16</sup> is not known. These three studies suggest that different loci might control resistance to different species of *Stagonosporopsis* causing GSB. Ideally, repeated studies using the same phenotyping conditions and rating systems should be conducted under field conditions, where evidence is emerging that populations of multiple species of *Stagonosporopsis* can exist within a single field<sup>37</sup>.

KASP Assay	Ta (°C)	SNP Position	Primer type	Primer sequence (5'–3')	Allele
ClGSB5.2-1	57	UGA5_25968975	FAM	GAAGGTGACCAAGTTCATGCT CCATTGGAATTG CCACTAGTTTGC	PI 189225 (R)
			VIC	GAAGGTGCGGAGTCAACGGATTCCATTGGAATTG CCACTAGTTTGT	SB (S)
			Reverse	TCAAAGGTCTGCTGGCTCCT	
ClGSB5.2-2	57	UGA5_26536280	FAM	GAAGGTGCGGAGTCAACGGATTCTACTTTCCAT GCACATGCTCTC	PI 189225
			VIC	GAAGGTGACCAAGTTCATGCTCTACTTTCCAT GCACATGCTCTA	SB
			Reverse	CCGGGTAAGTTCAGATCG	
ClGSB9.1-1	55	UGA9_8886865	FAM	GAAGGTGCGGAGTCAACGGATT AATGCATCAATA TCAAGCCAAATCTAA	PI 189225
			VIC	GAAGGTGACCAAGTTCATGCT AATGCATCAATA TCAAGCCAAATCTAG	SB
			Reverse	GTGAGATATTGTTTCTATGTGGAGCTAATA	
ClGSB11.1-1	55	UGA11_30871000	FAM	GAAGGTGCGGAGTCAACGGATT GTATGCCCTCTT AATTTGAGATTCTG	PI 189225
			VIC	GAAGGTGACCAAGTTCATGCT TGTATGCCCTCT TAATTTGAGATTCTA	SB
			Reverse	ATTCTTTGTCCGTTTTAGGTTGTACATT	

**Table 3.** KASP assays for flanking SNPs for GSB resistance QTL on chromosome 5 (*Qgsb5.2*).

*Qgsb5.2* is at a different location than the QTL identified by Gimode et al.<sup>17</sup> using the same *S. citrulli* isolate (12178A), but a different resistant parent (PI 482276). This seems to indicate that the two PIs contain distinct resistance loci. This could be beneficial to breeders as it would allow their effects to be combined through gene pyramiding.

Of the 38 predicted genes within the *Qgsb5.2* region, there are three genes of particular interest for their potential association with disease resistance. *Cl97C05G096980* is homologous to the histidine kinase 5 gene of *A. thaliana*, which has been shown to regulate the production of reactive oxygen species in response to stressors including necrotrophic fungi<sup>25</sup>. *Cl97C05G097030* is similar to the *enhanced disease resistance 2* gene in *A. thaliana* which has been shown to negatively regulate salicylic acid-based defenses and cell death in powdery mildew infections<sup>27</sup>. *Cl97C05G096890* is homologous to the *Ethylene-responsive transcription factor 2* gene from *A. thaliana*, which has been shown to be related to signaling pathways involved in resistance to *Alternaria brassicicola*<sup>38</sup>. Interestingly, salicylic acid-based signaling tends to be related to defense against biotrophic disease such as powdery mildew, while ethylene based signaling tends to be related to defense against necrotrophic diseases like those caused by *A. brassicicola*<sup>39</sup>. Given that GSB causing *Stagonosporopsis* species are thought to feed necrotrophically, *Cl97C05G096980* and *Cl97C05G096890* are likely candidate genes for this QTL. Further gene expression and gene knockout studies are needed to confirm the potential role of these genes in GSB resistance.

The current study used the *C. lanatus* 97103\_v2 genome<sup>24</sup> for analysis, since at the time, the *C. amarus* genome was not publicly available. However, additional resistance genes might be present in *C. amarus* that could be missed in the present study. It will be interesting to investigate these regions of interest further using the *C. amarus* genome and the watermelon Pan-genome that is currently in development (Shan Wu, personal communication).

The identification of *Qgsb5.2* brings the total number of QTL identified for GSB resistance to five. However, it remains unclear if these QTL contribute resistance to all GSB causing *Stagonosporopsis* species and isolates. More research is needed to determine the utility of these QTL against different species and isolates of *Stagonosporopsis* under field conditions. However, in light of the discovery of distinct QTL by this study, Ren et al.<sup>15</sup>, and Lee et al.<sup>17</sup> in germplasm derived from the same resistant PI, it is beginning to appear as if field level GSB resistance will require the incorporation of a number of different QTL<sup>15,16</sup>. Pyramiding multiple resistance genes can allow breeders to develop plants that are resistant to many different isolates or species of disease-causing pathogens, while increasing the durability of resistance to individual pathogen isolates, as shown for wheat stem rust<sup>40</sup>. QTL mapping resistance to multiple isolates of GSB-causing fungi in the same study would contribute greatly to our understanding of this issue and may provide the requisite QTL for pyramiding. The QTL identified in this study represents one part of that potential resistance gene pyramid to reduce the need to spray fungicides in order to control the disease.

## Materials and methods

**Plant materials.** A cross was made between Sugar Baby (*C. lanatus*, susceptible) and PI 189225 (*C. amarus*, resistant) and a single plant from the resulting F<sub>1</sub> was selfed to generate F<sub>2</sub> seed. F<sub>2</sub> plants were selfed in the greenhouse to produce 111 F<sub>2,3</sub> families. Leaf material collected from the parents, F<sub>1</sub> and F<sub>2</sub> plants were stored at –80 °C until use.

**Inoculum preparation.** *Stagonosporopsis citrulli* isolate 12178A (provided by M. Brewer, collected in Berrien County, Georgia, U.S.A in 2012) was grown on full strength potato dextrose agar (PDA) for four weeks and

then subcultured on quarter-strength PDA for a further 17 days. Conidia was harvested by flooding each Petri dish with 10 mL 0.1% Tween-20 (MilliporeSigma, St. Louis, MO) and scraping the plate with a microscope slide. The solution was filtered through 3 layers of cheese cloth and conidia were quantified under a microscope using a hemocytometer (Hausser Scientific, Horsham, PA). The conidial solution was diluted to 500,000 conidia per mL prior to inoculation. Due to low yields, a concentration of 420,000 conidia per mL was used for the re-screen.

**Resistance screening.** Initial resistance screening was conducted between March 20th and April 8th, 2019, in a greenhouse in Athens, GA. Seed from the parents,  $F_1$  and 111  $F_{2,3}$  families were sown in 48-well seedling trays in a randomized complete block design with four plants per block and four blocks. Two additional trays containing controls and 12 of the  $F_{2,3}$  families were also sown for mock inoculations. Seedlings were grown in a greenhouse with supplemental lighting until the 2–3 true leaf stage (16 days) and inoculated by spraying inoculum on the seedlings with a hand spray bottle until run-off. The mock inoculations were sprayed with 0.1% Tween-20 solution. After inoculation, the plants were placed inside a plastic tunnel with two humidifiers (Trion IAQ, Sanford, NC) in the greenhouse (average temperature = 24.0 °C, average humidity = 98.8%). After 72 h, the seedlings were moved to a greenhouse bench and overhead watered twice per day for four days. At 7 days post inoculation, the seedlings were rated for disease severity using a 0–10 scale (0: no lesions present on the first two true leaves, 1: 1–10% of first two true leaves covered in lesions, 2: 11–20% of first two true leaves covered in lesions, 3: 21–30% of first two true leaves covered in lesions, 4: 31–40% of first two true leaves covered in lesions, one leaf beginning to collapse, 5: 40–50% of first two true leaves covered in lesions, 1 true leaf collapsed, 6: 1 of first two true leaves dead, other 1–25% of true leaves covered in lesions, 7: 1 of first two true leaves dead, other 26–50% covered in lesions, 8: 1 of first two true leaves dead, other 51–75% covered in lesions, 9: both first two true leaves dead, 10: seedling totally collapsed or dead). The mean severity rating for each family/control was calculated using JMP® version 15.0.0 (SAS Institute Inc., Cary, NC). A Shapiro–Wilk test was conducted using JMP to test the normality of the phenotypic distribution. Broad sense heritability ( $H^2$ ) was calculated as described by Kearsley and Pooni<sup>41</sup>:  $H^2 = \frac{V_{F_2} - \frac{1}{4}(V_{P_1} + V_{P_2} + 2V_{F_1})}{V_{F_2}}$ , where  $V_{P_1}$ ,  $V_{P_2}$ ,  $V_{F_1}$  and  $V_{F_2}$  represents the variation among plants from parent 1, parent 2, the  $F_1$  and  $F_2$ , respectively.

The 20 most resistant and 20 most susceptible families were re-screened using the same protocol between May 1st and May 25th, 2019 (average temperature = 23.4 °C, average humidity = 91.7%). The 12 most resistant and most susceptible families over both screens were used to construct the bulks for QTL-seq. All methods were performed in accordance with the institutional and national Biosafety regulations.

**DNA extraction and sequencing.** Frozen leaf material from the  $F_2$  plants of the 24 selected  $F_{2,3}$  families was ground using a TissueLyser II (QIAGEN, Hilden, Germany) and DNA was extracted using the E.Z.N.A. HP Plant DNA Mini Kit (Omega Bio-Tek Inc., Norcross, GA). DNA was quantified with an Infinite M200 Pro plate reader (Tecan Group Ltd., Mannedorf, Switzerland) using the i-control software (Tecan Group Ltd.) and a NanoQuant Plate™ (Tecan Group Ltd.). Agarose gel electrophoresis was used to confirm the quality of genomic DNA. Equal amounts of DNA from the 12 most resistant and 12 most susceptible families were pooled to create the resistant (R-Bulk) and susceptible bulk (S-Bulk), respectively. The samples were sent to Novogene (Novogene Corporation Inc., Davis, CA) for whole genome sequencing on an Illumina Platform (PE150, Q30 > 80%).

**QTL-seq.** The raw reads were combined into forward and reverse read files for each bulk and checked for quality using FastQC<sup>42</sup>. The reads were then aligned to the *C. lanatus* 97103\_v2 genome using BWA-MEM<sup>24,43</sup>. The resulting SAM files were converted into BAM files using SAMtools, which was also used to sort the reads by alignment position and then index the resulting BAM files<sup>44</sup>. SAMtools was used to calculate mapping and pairing ratios of raw reads. BEDtools was then used to calculate average read depth for the BAM files<sup>45</sup>. Read group names were standardized using Picard Tools (<http://broadinstitute.github.io/picard/>), which was then used to mark duplicate reads and index the files. SAMtools was used to index the 97103\_v2 genome, and Picard Tools was used to create a dictionary file for the 97103\_v2 genome. Reads were re-aligned using GATK to generate clean reads from misaligned regions<sup>46</sup>. GATK was used to perform variant calling and the resulting VCF files from each bulk were combined into a single VCF file. GATK was used to filter the VCF file so that only SNPs remained. These SNPs were then filtered using GATK (QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0), and filtered SNPs were then removed using VCFtools<sup>47</sup>. GATK was used to output the SNPs from the VCF file into a table format suitable for use in the R package QTLseqr<sup>48,49</sup>. After importation into QTLseqr, SNPs were again filtered (refAlleleFreq = 0.2, minTotalDepth = 100, maxTotalDepth = 150, minSampleDepth = 40, minGQ = 99), and then used to calculate the  $\Delta$ -SNP index at each SNP<sup>18</sup>. This  $\Delta$ -SNP index was used with a 1 Mb sliding window to calculate a smoothed  $\Delta$ -SNP index as well as 95% and 99% confidence intervals for a region's contribution to the trait of interest (GSB resistance).

**Primer design and genotyping.** KASP (LGC Genomics LLC, Teddington, UK) primers were designed for SNPs spanning the regions of interest identified by QTLseqr using Primer3Plus and the *C. lanatus* 97103\_v2 genome<sup>24</sup>. DNA was extracted for parents,  $F_1$ , and all  $F_2$  plants as previously described. KASP PCR reactions contained 1.94  $\mu$ L of 2 × KASP Master Mix (LGC Genomics LLC), 0.06  $\mu$ L of KASP Primer Mix, and 2  $\mu$ L of DNA (10–20 ng/ $\mu$ L) in a total volume of 4  $\mu$ L. The KASP Primer Mix contained 12  $\mu$ L of each forward primer (100  $\mu$ M), 30  $\mu$ L of reverse primer (100  $\mu$ M), and 46  $\mu$ L of sterile distilled water. The following PCR conditions were used: 95 °C for 15 min followed by 10 touchdown cycles of 95 °C for 20 s, 66 °C for 25 s, and 72 °C for 15 s, followed by 35 cycles of 95 °C for 10 s, 57 °C for 60 s and 72 °C for 15 s. KASP fluorescent end readings were

measured using an Infinite M200 Pro (Tecan Group Ltd.) plate reader using the Magellan (Tecan Group Ltd.) software. Genotypes were called using KlusterCaller™ (LGC Genomics LLC).

**QTL mapping.** A genetic map (n = 111) was created of each region of interest using the following settings in ICIMapping<sup>50</sup>: Grouping was performed by recombination frequency with a threshold value of 0.30, ordering was performed k-Optimally by recombination frequency using the 2-OptMAP algorithm with 10 NN initials, rippling was performed by recombination frequency with a window size of 5. Resulting genetic maps were visualized using MapChart<sup>51</sup>. The genetic map was then used for QTL mapping using ICIMapping with the following settings<sup>50</sup>: The ICIM-ADD mapping method was used with deletion of missing phenotypes, a 1 cM step, and a value of 0.001 for the probability in stepwise regression. The LOD threshold of 2.5079 was determined by running 1000 permutations. The resulting data was graphed using R and the ggplot package<sup>49,52</sup>. Digenic epistasis was investigated using the ICIM-EPI method in ICIMapping<sup>50</sup>. One-way ANOVAs and Tukey–Kramer tests were performed using JMP® version 15.0.0 (SAS Institute Inc., Cary, NC) to test the association between markers and phenotype, as well as a haplotype representing flanking markers for *Qgsb5.2* simultaneously, and GSB disease severity in the population.

**Syntenic analysis and candidate genes.** The cucurbit genomics genome browser (<http://cucurbitgenomics.org>) was used to identify gene predictions for the *C. lanatus* 97103v2 genome<sup>24,53</sup>. These genes were then examined for predicted functions related to disease resistance. The cucurbit genomics syntenic browser (<http://cucurbitgenomics.org>) was used to check the syntenic of the *Qgsb5.2* to other reported GSB resistance QTL in cucurbits.

### Data availability

The sequencing data used during the current study available from NCBI Sequence Read Archive (PRJNA873687).

Received: 6 March 2022; Accepted: 3 November 2022

Published online: 15 November 2022

### References

- Keinath, A. P. From native plants in central Europe to cultivated crops worldwide: The emergence of *Didymella bryoniae* as a cucurbit pathogen. *HortScience* **46**, 532–535 (2011).
- Stewart, J. E., Turner, A. N. & Brewer, M. T. Evolutionary history and variation in host range of three *Stagonosporopsis* species causing gummy stem blight of cucurbits. *Fungal Biol.* **119**, 370–382 (2015).
- Keinath, A. P. & Duthie, J. A. Yield and quality reductions in watermelon due to anthracnose, gummy stem blight, and black rot. *Recent Research Developments in Plant Pathology* 77–90 (1998).
- Schwartz, H. F. & Gent, D. H. Gummy stem blight and black rot (cucumber, melon, pumpkin, squash, and zucchini). In: *High Plains IPM Guide* (Cooperative effort of the University of Wyoming, University of Nebraska, Colorado State University and Montana State University, 2007).
- Wightwick, A., Walters, R., Allinson, G., Reichman, S. & Menzies, N. Environmental risks of fungicides used in horticultural production systems. *Fungicides* **1**, 273–304 (2010).
- Norton, J. D. Inheritance of resistance to gummy stem blight [caused by *Didymella bryoniae*] in watermelon. *HortScience (USA)* **14**, 630–632 (1979).
- Norton, J. D., Cosper, R. D., Smith, D. A. & Rymal, K. S. ‘AU-Jubilant’ and ‘AU-Producer’ watermelons. *HortScience (USA)* **21**, 1460–1461 (1986).
- Norton, J., Boyhan, G. E., Smith, D. A. & Abrahams, B. R. AU-sweet scarlet: A disease resistant high quality watermelon with dark red flesh. (1993).
- Norton, J. D., Boyhan, G. E., Smith, D. A. & Abrahams, B. R. ‘AU-sweet scarlet’ watermelon. *HortScience* **30**, 393–394 (1995).
- Sumner, D. & Hall, M. Resistance of watermelon cultivars to Fusarium wilt and gummy stem blight. *Biol. Cult. Tests* **8**, 36 (1993).
- Boyhan, G., Norton, J. & Abrahams, B. Screening for resistance to anthracnose (race 2), gummy stem blight, and root knot nematode in watermelon germplasm. *Cucurbit Genet. Coop. Rep.* **17**, 106–110 (1994).
- Song, R., Gusmini, G. & Wehner, T. C. Screening the watermelon germplasm collection for resistance to gummy stem blight. *XXVI International Horticultural Congress: Advances in Vegetable Breeding Acta Horticulturae* 63–68 (2004).
- Gusmini, G., Song, R. & Wehner, T. C. New sources of resistance to gummy stem blight in watermelon. *Crop Sci.* **45**, 582–588 (2005).
- Gusmini, G., Rivera-Burgos, L. A. & Wehner, T. C. Inheritance of resistance to gummy stem blight in watermelon. *HortScience* **52**, 1477–1482 (2017).
- Ren, R. *et al.* Identification and molecular mapping of a gummy stem blight resistance gene in wild watermelon (*Citrullus amarus*) germplasm PI 189225. *Plant Dis.* **104**, 16–24 (2019).
- Lee, E. S. *et al.* QTL mapping for gummy stem blight resistance in watermelon (*Citrullus* spp.). *Plants* **10**, 500 (2021).
- Gimode, W., Bao, K., Fei, Z. & McGregor, C. QTL associated with gummy stem blight resistance in watermelon. *Theor. Appl. Genet.* **134**, 573–584 (2021).
- Takagi, H. *et al.* Qtl-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **74**, 174–183 (2013).
- Michelmore, R. W., Paran, I. & Kesseli, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *PNAS* **88**, 9828–9832 (1991).
- Clevenger, J. *et al.* Mapping late leaf spot resistance in peanut (*Arachis hypogaea*) using QTL-seq reveals markers for marker-assisted selection. *Front. Plant Sci.* **9**, 83 (2018).
- Branham, S. E. *et al.* QTL-seq and marker development for resistance to *Fusarium oxysporum* f. sp. *niveum* race 1 in cultivated watermelon. *Mol. Breed.* **38**, 139 (2018).
- Fall, L. A., Clevenger, J. & McGregor, C. Assay development and marker validation for marker assisted selection of *Fusarium oxysporum* f. sp. *niveum* race 1 in watermelon. *Mol. Breed.* **38**, 130 (2018).
- Sowell, G. & Pointer, G. R. Gummy stem blight resistance of introduced watermelons. (1962).
- Guo, S. *et al.* Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits. *Nat. Genet.* **51**, 1616–1623 (2019).



25. Pham, J. & Desikan, R. Modulation of ROS production and hormone levels by AHK5 during abiotic and biotic stress signaling. *Plant Signal. Behav.* **7**, 893–897 (2012).
26. Brown, R. L., Kazan, K., McGrath, K. C., Maclean, D. J. & Manners, J. M. A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of Arabidopsis. *Plant Physiol.* **132**, 1020–1032 (2003).
27. Vorwerk, S. *et al.* EDR2 negatively regulates salicylic acid-based defenses and cell death during powdery mildew infections of Arabidopsis thaliana. *BMC Plant Biol.* **7**, 35 (2007).
28. Lou, L. *et al.* Genetic mapping of gummy stem blight (*Didymella bryoniae*) resistance genes in *Cucumis sativus-hystrix* introgression lines. *Euphytica* **192**, 359–369 (2013).
29. Liu, S. *et al.* Genetic analysis and QTL mapping of resistance to gummy stem blight in *Cucumis sativus* seedling stage. *Plant Dis.* **101**, 1145–1152 (2017).
30. Zhang, S. *et al.* Inheritance and QTL mapping of resistance to gummy stem blight in cucumber stem. *Mol. Breed.* **37**, 49 (2017).
31. Gimode, W., Lonnee, M. & McGregor, C. Resistance response of *Citrullus* genotypes to *Stagonosporopsis* spp. Isolates causing gummy stem blight. *Cucurbit Genet. Coop. Rep.* **42**, 1–6 (2019).
32. Hassan, M. Z. *et al.* Genome-wide characterization of NBS-encoding genes in watermelon and their potential association with gummy stem blight resistance. *Int. J. Mol. Sci.* **20**, 902 (2019).
33. Zuniga, T. L., Jantz, J. P., Zitter, T. A. & Jahn, M. K. Monogenic dominant resistance to gummy stem blight in two melon (*Cucumis melo*) accessions. *Plant Dis.* **83**, 1105–1107 (1999).
34. Frantz, J. D. & Jahn, M. M. Five independent loci each control monogenic resistance to gummy stem blight in melon (*Cucumis melo* L.). *Theor. Appl. Genet.* **108**, 1033–1038 (2004).
35. Ramos, A., Fu, Y., Michael, V. & Meru, G. QTL-seq for identification of loci associated with resistance to *Phytophthora* crown rot in squash. *Sci. Rep.* **10**, 5326 (2020).
36. Vales, M. *et al.* Effect of population size on the estimation of QTL: A test using resistance to barley stripe rust. *Theor. Appl. Genet.* **111**, 1260–1270 (2005).
37. Li, H.-X. & Brewer, M. T. Spatial genetic structure and population dynamics of gummy stem blight fungi within and among watermelon fields in the southeastern United States. *Phytopathology* **106**, 900–908 (2016).
38. McGrath, K. C. *et al.* Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. *Plant Physiol.* **139**, 949–959 (2005).
39. Spoel, S. H., Johnson, J. S. & Dong, X. Regulation of tradeoffs between plant defenses against pathogens with different lifestyles. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18842–18847 (2007).
40. Zhang, B. *et al.* Pyramiding stem rust resistance genes to race TTKSK (Ug99) in wheat. *Can. J. Plant Path.* **41**, 443–449 (2019).
41. Kearsey, M. J. & Pooni, H. S. *The genetical analysis of quantitative traits* (Chapman and Hall, 1996).
42. Andrews, S. In: *FastQC: a quality control tool for high throughput sequence data* (Babraham Bioinformatics Babraham Institute, 2010).
43. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997* (2013).
44. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
45. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
46. McKenna, A. *et al.* The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
47. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
48. Mansfeld, B. N. & Grumet, R. QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. *Plant Genome* **11**, 180006 (2018).
49. R Core Team, Rf. & others. R: A language and environment for statistical computing. R foundation for statistical computing Vienna, Austria, (2013).
50. Meng, L., Li, H., Zhang, L. & Wang, J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* **3**, 269–283 (2015).
51. Voorrips, R. E. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **93**, 77–78 (2002).
52. Wickham, H. *ggplot2: elegant graphics for data analysis* (Springer-Verlag, 2016).
53. Zheng, Y. *et al.* Cucurbit genomics database (CuGenDB): A central portal for comparative and functional genomics of cucurbit crops. *Nucleic Acids Res.* **47**, D1128–D1136 (2019).

## Acknowledgements

This research was partially funded by the National Institute of Food and Agriculture, U.S. Department of Agriculture, award numbers 2015-51181-24285 and 2020-51181-32139. We thank Winnie Gimode for her assistance with phenotyping, and Manoj Sapkota and Yasin Topcu for their assistance with QTLseq and data visualization.

## Author contributions

L.F.A. conducted all experiments, data collection and analysis and wrote the manuscript. C.M. conceived the project, provided guidance during the project and edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022