Characterization and QTL Mapping of Parthenocarpic Fruit Set in Processing Cucumber (*Cucumis sativus* L.)

by

Calvin Dean Lietzow

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

(Plant Breeding and Plant Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2014

Date of final oral examination: 06/06/2014

The dissertation is approved by the following members of the Final Oral Committee:

Yiqun Weng, Associate Professor, Horticulture
Michael Havey, Professor, Horticulture
Philipp Simon, Professor, Horticulture
William Tracy, Professor, Agronomy
Juan Zalapa, Assistant Professor, Horticulture

Characterization and QTL Mapping of Parthenocarpic Fruit Set in Processing Cucumber (*Cucumis sativus* L.)

Calvin Dean Lietzow

Under the supervision of Professor Yiqun Weng

at the University of Wisconsin-Madison

Abstract

Parthenocarpy is a desirable trait for cucumber production and is particularly valuable in environments where pollination is difficult or adversely affected by abiotic factors. Parthenocarpic cucumber cultivars have been successfully developed, but the genetic and molecular mechanisms behind parthenocarpic expression remain largely unknown. Since parthenocarpy is often considered a yield component, it is difficult to separate the true parthenocarpic character from other yield related traits. Therefore, this study was designed to better define what is considered true parthenocarpic expression and then to use this knowledge to identify QTL associated with parthenocarpic fruit set. Building off of previous studies demonstrating that parthenocarpic fruit set is initiated in the days before and immediately after anthesis, a new approach to phenotyping parthenocarpic fruit set in cucumber focused on early fruit initiation and development was implemented. With a clear approach to phenotypic evaluation, a mapping population consisting of 205 F_3 families derived from a cross between processing cucumber inbred lines '2A' (parthenocarpic) and 'Gy8' (non-parthenocarpic) was evaluated for parthenocarpic fruit set. Genotypic data collected for each F_2 individual was utilized to construct a linkage map consisting of 192 marker loci in seven linkage groups and covering 571.7 cM. Multiple QTL mapping methodologies (interval mapping, composite

interval mapping, and multiple interval mapping) were employed to detect and construct optimal models for the inheritance of parthenocarpic fruit set. Seven additive QTL associated with parthenocarpic fruit set were detected with four identified consistently in all analyses. The four consensus QTL were located on chromosome 5 at 32.3 - 54.7 cM, chromosome 6 at 0.0 - 9.7 cM, chromosome 6 at 80.0 - 83.0 cM, and chromosome 7 at 21.8 - 32.1 cM. Bioinformatic analysis of the genomic regions harboring the four consensus QTL was conducted and multiple candidate genes were identified. A model was proposed to explain the roles of potential candidate genes in parthenocarpic expression observed in cucumber. The QTL identified for parthenocarpic fruit set by this study are valuable to cucumber breeders interested in developing parthenocarpic cultivars and to researchers interested in the genetic and molecular mechanisms of parthenocarpic fruit set.

Acknowledgements

I would like to thank my mentor and friend, Dr. Yiqun Weng, for providing me with encouragement, patience, and the opportunity to pursue a graduate degree. I would like to thank Dr. Michael Havey, Dr. Philipp Simon, Dr. William Tracy, and Dr. Juan Zalapa for their advice and encouragement while serving as my thesis committee. I would like to thank all of my fellow Weng Lab colleagues from the past few years for their eagerness to help with even the smallest tasks. Finally, I would like to thank all of my family and friends. I would not have been able to reach this accomplishment without their unending support and guidance.

Table of Contents

Abstract i

Acknowledgements iii

Table of Contents iv

List of Tables vi

List of Figures ix

List of Addendums xii

Introduction 1

Chapter 1: A Novel Approach to Phenotypic Evaluation of Parthenocarpic Fruit Set in Processing Cucumber (*Cucumis sativus* L.)

- Abstract 16
- Introduction 17
- Materials and Methods 19
- Results and Discussion 24
- Literature Cited 31

Chapter 2: Construction of a Linkage Map in an $F_{2:3}$ Population Segregating for Parthenocarpic Fruit Set in Cucumber (*Cucumis sativus* L.)

- Abstract 62
- Introduction 63
- Materials and Methods 66
- Results and discussion 70
- Literature Cited 75

Chapter 3: Identification of Quantitative Trait Loci Associated with Parthenocarpic Fruit Set in Processing Cucumber (*Cucumis sativus* L.)

Abstract 102

Introduction 103

Materials and Methods 106

Results and Discussion 113

Literature Cited 135

List of Tables

Chapter 1

Table 1. The number of fruit initiated per plant in each of the three greenhouse experiments conducted to study parthenocarpic fruit set in a 2A×GY8 $F_{2:3}$ population of *C. sativus* 36

Table 2. The p-values obtained from two sample t-tests used for comparing of each of the three greenhouse experiments conducted to study parthenocarpic fruit set in a 2A×Gy8 $F_{2:3}$ population of *C. sativus* 37

Table 3. Spearman's rank correlation coefficients (Spearman, 1904) from comparisons of F_3 family means in each of the three greenhouse experiments conducted to study parthenocarpic fruit set in a 2A×Gy8 $F_{2:3}$ population of *C. sativus*. Due to changes in experimental design between experiment 1 and experiments 2 and 3, which were conducted identically, rho values are also presented comparing experiment 1 with a data set consisting of combined data from experiments 2 and 3 38

Table 4. Comparisons of the number of fruit initiated on plants of inbred lines '2A' and 'Gy8' under different plant maintenance regimens. Pruned plants were cleared of all lateral branches, foliage, and flowers from the bottom five nodes of the plant. Unpruned plants were not pruned during growth and allowed to grow unhindered. The Unpruned With Laterals and Unpruned datasets utilize the same plants but are evaluated under two different protocols. The Pruned dataset is collected from separate plants.....39

Chapter 2

Table 1. General statistics obtained for the linkage map generated for a 2A×Gy8 $F_{2:3}$ population developed for the study of parthenocarpic fruit set in *C. sativus*. The linkage map was constructed with 205 2A×Gy8 F_2 individuals and consists of 185 SSR, 5 STS, and 2 dCAPS marker loci in seven linkage groups covering 571.7 cM. Linkage groups were assigned to chromosomes according to Yang et al. (2012). Map distances were calculated using Kosambi's map function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001) 81

Table 2. The number of SSR primer pairs showing polymorphisms between the parthenocarpic parent inbred line '2A' and the non-parthenocarpic parent inbred line 'Gy8' in 9% denaturing polyacrylamide gel electrophoresis (PAGE) 82

Table 3. Table presenting the segregation data for each marker locus contained in the 2AxGy8 linkage map. The linkage map consists of 185 SSR, 5 STS, and 2 dCAPs marker loci in seven

linkage groups. The "A" genotype has been assigned to '2A' while the "B" genotype represents 'Gy8'. Distortion from the expected 1:2:1 segregation for co-dominant markers is evaluated by the chi square test. The linkage map contains two SSR markers on Chromosome 6, UW084474 and UW026722, that could only be visually scored as dominant markers with the protocol outlined by this study. These markers were evaluated with an expected segregation ratio of 3:1 83

Table 4. Identification and location of intervals exceeding 3 Mb in the linkage map without marker coverage. The linkage map was constructed for a *C. sativus* $F_{2:3}$ population consisting of 205 F_2 individuals derived from a cross of parthenocarpic inbred line '2A' and non-parthenocarpic inbred line 'Gy8'. All SSR, STS, and dCAPS markers screened for polymorphisms between the parental lines were analyzed with *in silico* PCR using the Gy14 draft genome assembly version 1.0 (Yang et al., 2012). The table presents the number of markers identified by *in silico* PCR as being located within the intervals unaccounted for by the linkage map. None of the markers in these intervals were found to be polymorphic in the 2A×Gy8 population. It should be noted that *in silico* PCR was only able to confirm PCR amplicons for 77.6% of markers screened by this study. Consequently, the number of markers screened within these intervals is likely greater, but the data presented provides an estimate of marker availability in these regions 89

Table 5. Scaffold and assembly position data obtained from Gy14 and 9930 draft genome assemblies for the linkage map generated for a 2A×Gy8 $F_{2:3}$ population developed for the study of parthenocarpic fruit set in *C. sativus* (Huang et al., 2009; Yang et al., 2012). Genome assembly position data is only presented for 'Gy14' for simplicity, as genome assembly positions for '9930' are similar in order.....90

Chapter 3

Table 1. QTL mapping results for parthenocarpic fruit set in the pooled data obtained from experiments 1-3. The pooled data consists of 205 F_3 families represented by 11 individuals each. Phenotypic data was collected from nodes 5-30 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).....143

Table 2. QTL mapping results for parthenocarpic fruit set in the data obtained from experiment 1. The data consists of 201 F_3 families represented by five individuals each. Phenotypic data was collected from nodes 5-30 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003) 145

Table 3. QTL mapping results for parthenocarpic fruit set in the dataset obtained from the compilation of experiments 2 and 3. The data consists of 205 F_3 families represented by six

individuals each. Phenotypic data was collected from nodes 5-30 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).....147

Table 4. QTL mapping results for parthenocarpic fruit set in the dataset obtained from the compilation of experiments 2 and 3. The data consists of 205 F_3 families represented by six individuals each. Phenotypic data was collected from nodes 5-20 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).....149

Table 5. QTL mapping results for parthenocarpic fruit set in the dataset obtained from the compilation of experiments 2 and 3. The pooled data consists of 205 F_3 families represented by six individuals each. Phenotypic data was collected from nodes 5-10 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003) 151

Table 6. QTL mapping results for seed size. Data was collected from F_3 seed obtained by self-pollination of 205 2A×Gy8 F_2 plants. Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds. All QTL analyses were performed with R/qtl software (Broman et al., 2003). 152

Table 7. QTL mapping results for seed weight. Data was collected from F_3 seed obtained by self-pollination of 205 2A×Gy8 F_2 plants. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit. All QTL analyses were performed with R/qtl software (Broman et al., 2003) 153

List of Figures

Chapter 1

Figure 1. A: Photograph illustrating the differences between dried unpollinated *C. sativus* ovaries that have initiated parthenocarpic fruit development prior to aborting (left) and dried unpollinated ovaries that aborted at anthesis (right). At center are unpollinated ovaries at anthesis for comparison. B: Photograph illustrating examples of parthenocarpic fruit that have aborted after multiple days of fruit development.....40

Figure 2. Frequency distributions of the total number of fruit initiated per plant in experiments 1-3. Data was collected from plant nodes 6 thru 30 for a maximum possible total of 25 initiated fruit per plant. In each experiment the average values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Experiment 1 consisted of 1050 2A×Gy8 F_3 and accompanying control plants. B: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants. C: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants 41

Figure 3. Frequency distributions of the total number of fruit initiated per plant in the pooling of experiments 1-3. Data was collected from plant nodes 6 thru 30 for a maximum possible total of 25 initiated fruit per plant. In each figure the average pooled values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Frequency distribution of the total number of fruit initiated per plant in the pooling of experiments 1-3. The pooled experiments together consisted of 2370 2A×Gy8 F_3 and accompanying control plants. B: Frequency distribution of the average number of fruit initiated for 205 F_3 families obtained by pooling across the three experiments. Each F_3 family is represented by 11 F_3 plants.....42

Figure 4. Frequency distributions displaying the frequency of fruit set initiation at each plant node across experiments 1-3. All plants were cleared of flowers and vegetation on nodes 1 thru 5. A: Experiment 1 consisted of 1050 2A×Gy8 F_3 and accompanying control plants. B: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants. C: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants 43

Figure 5. Frequency distributions of the total number of fruit initiated per plant in experiments 2 and 3. Data was collected from plant nodes 6 thru 20 for a maximum possible total of 15 initiated fruit per plant. In each experiment the average values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants and was grown from September 2011 to December 2011. B: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants and was grown from March 2012 to June 2012.....44

Figure 6. Photograph depicting the typical fruit number and fruit set location in *C. sativus* parental lines '2A' (left) and 'Gy8' (right) and the 2A×Gy8 F_1 hybrid (center) 45

Figure 7. Frequency distributions of the total number of fruit initiated per plant in experiments 2 and 3. Data was collected from plant nodes 6 thru 10 for a maximum possible total of five initiated fruit per plant. In each experiment the average values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants. B: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants.....46

Figure 8. Photographs of '2A' (A) and 'Gy8' (B) under pruned and unpruned treatments in environmentally controlled greenhouse conditions. In each photograph the plant on the left was not trimmed of lateral branches or cleared of foliage from the bottom five nodes of the plant. The plant on the right was trimmed of all lateral branches routinely throughout growth and was cleared of all foliage from the bottom five nodes of the plant.....47

Chapter 2

Figure 1. Linkage map constructed for a $F_{2:3}$ mapping population that was generated from the cross of a highly parthenocarpic inbred line, '2A', with a non-parthenocarpic inbred line, 'Gy8'. The resulting linkage map was constructed with 205 2A×Gy8 F_2 individuals and consists of 185 SSR, 5 STS, and 2 dCAPS marker loci in seven linkage groups covering 571.7 cM. Map distances were calculated using Kosambi's map function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001). Linkage groups were assigned to chromosomes according to Yang et al. (2012).....98

Chapter 3

Figure 1. Plot of genome wide LOD curves obtained by interval mapping from data collected for experiments 1-3. After experiment 1 was observed to differ from experiments 2 and 3 in Chapter 1, a separate QTL analysis was conducted. Pooled data from experiments 1-3 is compared with data from experiment 1 and a dataset composed of experiments 2 and 3 combined. QTL analyses were performed with R/qtl software (Broman et al., 2003) 154

Figure 2. Plot of LOD curves for chromosomes 2(A), 4(B), 5(C), 6(D), and 7(E) obtained by interval mapping from data collected for experiments 1-3. After experiment 1 was observed to differ from experiments 2 and 3 in Chapter 1, a separate QTL analysis was conducted. Pooled data from experiments 1-3 is compared with data from experiment 1 and a dataset composed of experiments 2 and 3 combined. QTL analyses were performed with R/qtl software (Broman et al., 2003) 155

Figure 3. Plot of genome wide LOD curves obtained by interval mapping from data collected from the first 10, 20, and 30 nodes of each F_3 plant. For this comparison only data from experiments 2 and 3 were included due to the delayed fruit set observed in experiment 1. QTL analyses were performed with R/qtl software (Broman et al., 2003) 156

Figure 4. Plot of LOD curves for chromosomes 5(A), 6(B), and 7(C) obtained by interval mapping from data collected from the first 10, 20, and 30 nodes of each F_3 plant. For this comparison only data from experiments 2 and 3 were included due to the delayed fruit set observed in experiment 1. QTL analyses were performed with R/qtl software (Broman et al., 2003) 157

Figure 5. Plot of genome wide LOD curves obtained by interval mapping from data collected for the seed size of each F_2 plant. Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds. QTL analyses were performed with R/qtl software (Broman et al., 2003). 158

Figure 6. Plot of LOD curves for chromosomes 5(A) and 6(B) obtained by interval mapping from data collected for the seed size and weight of each F_2 plant. Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit. QTL analyses were performed with R/qtl software (Broman et al., 2003) 159

Figure 7. Plot of genome wide LOD curves obtained by interval mapping from data collected for the seed weight of each F_2 plant. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit. QTL analyses were performed with R/qtl software (Broman et al., 2003) 160

Figure 8. Plot of LOD curves for comparison of parthenocarpic fruit set and seed weight traits. Data for parthenocarpic fruit set is obtained from a data set consisting of data from experiments 2 and 3 combined. Parthenocarpic fruit set is measured as the number of parthenocarpic fruits initiating growth on the first 30 nodes of each F_3 plant. Data is presented as F_3 family means. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit of each F_2 plant. A: LOD curves of whole genome scans of both traits obtained by interval mapping. B and C: LOD curves obtained by interval mapping for both traits on individual chromosomes 5(B) and 6(C).....161

List of Addendums

Chapter 1

Addendum 1. Descriptions of the origin and trait characteristics of inbred lines '2A', 'Gy8', and 'Gy7' 48

Addendum 2. Schematic for the plant spacing and layout in each of the five greenhouses utilized in experiment 1 at the University of Wisconsin-Madison Walnut Street Research Greenhouses located in Madison, WI. Experiment 1 consisted of 1050 2A×Gy8 F_3 and accompanying control plants. Each greenhouse in experiment 1 contained 1 F_3 plant from each of 201 2A×Gy8 F_3 families and 9 control plants for a total of 210 plants.....50

Addendum 3. Schematic for the plant spacing and layout in each of the five greenhouses utilized in experiments 2 and 3 at the University of Wisconsin-Madison Walnut Street Research Greenhouses located in Madison, WI. Experiments 2 and 3 each consisted of 660 2A×Gy8 F_3 and accompanying control plants. In experiments 2 and 3, three plants from each of 205 2A×Gy8 F_3 families were randomly distributed across each experiment. Each greenhouse contained 123 randomized 2A×Gy8 F_3 plants and 9 control plants for a total of 132 plants. 51

Addendum 4. Heat map displaying the spatial position and the number of initiated fruits set on each experimental plant during experiment 1 in the study of parthenocarpic fruit set. Numbers 1-5 represent each of the five individual greenhouses utilized during experiment 1. The numbers listed inside each shaded box indicate the 2A×Gy8 F_3 family or control line membership of each plant. Yellow blocks indicate plants that were either damaged or failed to reach maturity. Parental line '2A' is denoted as "P1", while parental line 'Gy8' is denoted as "NP1".....52

Addendum 5. Heat map displaying the spatial position and the number of initiated fruits set on each experimental plant during experiment 2 in the study of parthenocarpic fruit set. Numbers 1-5 represent each of the five individual greenhouses utilized during experiment 2. The numbers listed inside each shaded box indicate the 2A×Gy8 F_3 family or control line membership of each plant. Yellow blocks indicate plants that were either damaged or failed to reach maturity. Parental line '2A' is denoted as "P1", while parental line 'Gy8' is denoted as "NP1".....53

Addendum 6. Heat map displaying the spatial position and the number of initiated fruits set on each experimental plant during experiment 3 in the study of parthenocarpic fruit set. Numbers 1-5 represent each of the five individual greenhouses utilized during experiment 3. The numbers listed inside each shaded box indicate the 2A×Gy8 F_3 family or control line membership of each plant. Yellow blocks indicate plants that were either damaged or failed to reach maturity. Parental line '2A' is denoted as "P1", while parental line 'Gy8' is denoted as "NP1".....54

Addendum 7. The maturity of non-parthenocarpic cucumber cultivars at the time of harvest in 2012 commercial production trials at the University of Wisconsin-Madison Agricultural Research Station in Hancock, WI. Maturity was measured by the number of nodes on the main stem. Ten plants from each of three commercial plots were sampled. Two of the plots were planted with the same 'Excursion' cultivar 55

Addendum 8. 2A×Gy8 F_3 family means for each experiment conducted to study parthenocarpic fruit set in cucumber 56

Chapter 2

Addendum 1. Photograph demonstrating the generation of SSR molecular marker data by size-fractionation of PCR amplicons using polyacrylamide gel electrophoresis (PAGE) and silver staining. The PCR amplicons in this example were amplified with the SSR marker SSR10783. Lane 1 contains PCR amplicons amplified from parthenocarpic parent inbred line '2A'. Lane 2 contains PCR amplicons amplified from non-parthenocarpic parent inbred line 'Gy8'. Lane 5 contains GeneRuler 100 bp PLUS DNA ladder. All other lanes contain PCR amplicons amplified from 2A×Gy8 F_2 individuals in numerical order. Data in this example was collected from the major polymorphic bands located below 100bp of the DNA ladder (lowest band of the DNA ladder).....99

Addendum 2. Linkage map constructed with genotypic data obtained from 89 of 126 2AxGy8 F_2 individuals presented in Sun et al. (2006). To anchor the linkage map and provide points for direct comparison to the Gy14 draft genome assembly and the 2AxGy8 linkage map constructed in this study, marker data was collected from an additional 43 SSR markers. The combined linkage map consists of 236 marker loci contained in eight linkage groups. The eighth linkage group is the result of a split of the linkage group corresponding to Chromosome 1. All chromosome labels starting in "S" are from the Sun et al. study and labels beginning in "L" are from the current study.....100

Chapter 3

Addendum 1. Photograph depicting the difference in seed size between the parental inbred lines'2A' (right) and 'Gy8' (left). The seeds shown in this photo are healthy and fully developed seeds and are representative of seeds from each parental inbred line.....162

Addendum 2. The locations of QTL identified by the current study for "parthenocarpic fruit set" are compared with the locations of QTL identified for "parthenocarpic yield" by Sun et al. (2006) and "number of fruits per plant" by Fazio et al. (2003). All QTL are placed onto the linkage map constructed for the 2A×Gy8 $F_{2:3}$ mapping population used in the current study. With the use of the common SSR markers placed onto the linkage map of Sun et al., the approximate locations of QTL are identified for the Sun et al. study. Common AFLP and RAPD

markers between the linkage maps constructed by Sun et al. and Fazio et al. allowed for the approximate placement of the QTL reported by Fazio et al. 163

Addendum 3. Close up view of the LOD curves within the 1.5 LOD confidence interval for each of the four consensus QTL associated with parthenocarpic fruit set on chromosomes 5 (A), 6 (B and C), and 7 (D). Colored dashes mark the approximate location of each candidate gene. Candidate gene locations were derived by a BLASTn search of the predicted candidate gene sequence to the Gy14 Draft Genome Assembly Version 1.0 (Yang et al., 2012). These positions were then compared to known molecular marker positions within the assembly. Molecular markers are marked as tick marks on the x-axis of each graph (refer to Chapter 2 Table 5 for marker names and physical position within the assembly).....164

Addendum 4. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth6.2* QTL (sequence between molecular markers SSR17604 and UWSTS0310) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.....165

Addendum 5. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth6.1* QTL (sequence between molecular markers UWSTS0316 and SSR19672) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.....169

Addendum 6. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth7.1* QTL (sequence between molecular markers SSR00015 and UW085407) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.....173

Addendum 7. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth5.1* QTL (sequence between molecular markers UW001903 and SSR13409) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.....177

Addendum 8. Phenotypic data used in all analyses of the 2A×Gy8 $F_{2:3}$ cucumber population. The seed size and seed weight traits were collected from the F_2 generation. Parthenocarpic fruit set data was collected from the F_3 generation and presented as a mean of those values for each F_3 family 181

Addendum 9. Correlation coefficients calculated from comparisons of parthenocarpic fruit set, seed size, and seed weight traits in a 2AxGy8 $F_{2:3}$ cucumber population.....187

Addendum 10. Alignment of the predicted protein sequences of the candidate gene BRI1, obtained from the parental lines '2A' and 'Gy8'. Protein sequences were predicted with assembled sequence data obtained from whole genome re-sequencing of the parental lines. The predicted BRI1 protein from 'Gy14' was constructed from sequence data extracted from the Gy14 Draft Genome Assembly Version 1.0 and is included as a reference (Yang et al., 2012). Protein prediction was performed with the FGENESH utility provided by Softberry (Solovyev et al., 2006). The gap in sequence data observed for 'Gy8' is due to a gap between contigs of the 'Gy8' assembled re-sequencing data. An asterisk marks a potential polymorphism between '2A' and the other sequences 188

Addendum 11. Alignment of the predicted protein sequences of the candidate gene BAK1, obtained from the parental lines '2A' and 'Gy8'. Protein sequences were predicted with assembled sequence data obtained from whole genome re-sequencing of the parental lines. The predicted BAK1 protein from 'Gy14' was constructed from sequence data extracted from the Gy14 Draft Genome Assembly Version 1.0 and is included as a reference (Yang et al., 2012). Protein prediction was performed with the FGENESH utility provided by Softberry (Solovyev et al., 2006). The gap in sequence data observed for '2A' and 'Gy8' is due to gaps between contigs of the assembled re-sequencing data. The mismatch of sequence flanking the gap between contigs of '2A' is a result of an overhanging base pair attached to the edge of the first contig which resulted in a change to the predicted protein around this gap. An asterisk marks a potential polymorphism between '2A' and the other sequences 190

Introduction

Fertilization and fruit development are critical to angiosperm reproduction and dispersal. Upon successful pollination and fertilization, a number of physiological events occur that lead to fruit set and the development of the fruit and seed. Fruit development can be divided into four major phases (Gillaspy et al., 1993). Phase 1 includes ovary development, fertilization, and fruit set. During phase 1, cell division is reduced until physiological cues associated with pollination are received, which determine whether to proceed with or to abort fruit set (Gillaspy et al., 1993). With successful fruit initiation, phase 2 is characterized by rapid cell division, which occurs for approximately 7-10 days (Gillaspy et al., 1993). Phase 3 consists of rapid cell expansion, which can see a fruit increase in size by a factor of 100 fold or more (Gillaspy et al., 1993; Coombe, 1976). Fruit development concludes with the onset of fruit ripening in phase 4.

Fruit set can occur independent of pollination and/or fertilization and is referred to as parthenocarpy. Parthenocarpic fruit set can be induced with the application of compatible foreign pollen or aqueous pollen extracts to the stigma (Fitting, 1909; Gustafson, 1937; Noll, 1902; Yasuda, 1930; Yasuda, 1934; Yasuda, 1935; Srivastava, 2002). Pollen and pollen extracts are known to contain auxins, gibberellins, and brassinosteroids, which among other phytohormones may trigger fruit set (Gustafson, 1937; Gustafson, 1942; Tsao, 1980; Srivastava, 2002). Indeed, the exogenous application of phytohormones has been widely utilized in the manipulation of parthenocarpic crops (Pandolfini et al., 2009; Schwabe and Mills, 1981). Parthenocarpic fruit set can be influenced by events occurring away from the ovary. For example, removal of plant apical meristems induces parthenocarpic fruit set; presumably through alterations of hormone signaling (Carbonell and Garcia-Martinez, 1980; Coombe, 1962; Parry,

1976; Quinlan and Preston, 1971; Saunders et al., 1991; Serrani et al., 2010; Westwood and Bjornstad, 1974). Efficacy of fruit set can also be influenced by abiotic factors such as temperature, humidity, and low light intensity (George et al., 1984; Picken, 1984; Pike and Peterson, 1969; Vardy et al., 1989a; Vardy et al., 1989).

Parthenocarpy can be invoked artificially or observed naturally. Genetic parthenocarpy can be either obligatory or facultative. Obligatory parthenocarpy defines instances when a plant can only produce parthenocarpic fruit. With facultative parthenocarpy, fruits develop parthenocarpically only in the absence of fertilization. If successful pollination and fertilization occur, fruit will develop with fertile seeds. Within obligatory and facultative parthenocarpy, a distinction is made between vegetative and stimulated parthenocarpy. Vegetative parthenocarpy occurs when fruit set is observed without pollination. Fruit set through stimulated parthenocarpy requires pollination but fertilization is prevented or fails to occur. Parthenocarpy is distinct from, but often confused with stenospermocarpy. Stenospermocarpy occurs when both pollination and fertilization occur but the embryo aborts shortly after. Fruits continue to develop without viable seeds although traces of the seed coats are often observed.

Naturally occurring parthenocarpy has a genetic basis. Tomato has served as model organism for fruit development studies and consequently the inheritance of parthenocarpy has been well characterized in the crop. Four independent recessive genes, *pat* (*parthenocarpic fruit*), *pat-2*, and the two gene *pat-3*/*pat-4* background, have each been identified as inducers of facultative parthenocarpy in tomato (Nuez et al., 1986; Philouze and Maisonneuve, 1978; Philouze, 1989; Philouze and Pecaut, 1986; Soressi and Salamini, 1975; Vardy et al., 1989a). The *pat-2* and *pat-3*/*pat-4* genes are associated with increased levels of bioactive gibberellins within ovaries before pollination (Fos et al., 2000; Fos et al., 2001). Parthenocarpic fruits

obtained with *pat* are often undersized and thus deemed undesirable in cultivar development (Falavigna et al., 1978; Philouze and Pecaut, 1986). The parthenocarpic fruits of lines expressing *pat-2* are normal size but *pat-2* has been found to be pleiotropic and results in reduced yield and vigor in some genetic backgrounds (Philouze et al., 1988). Parthenocarpic lines obtained with *pat-3*/*pat-4* have normal sized fruits. However, when seeded fruits are set on the same plant, the developing parthenocarpic fruits will be substantially smaller than the seeded fruit. This along with its polygenic inheritance again makes *pat-3*/*pat-4* less than ideal for breeding proposes (Gorguet et al., 2005; Philouze, 1989). In addition to the *pat* genes, Gorguet et al. (2008) identified three unique QTL associated with parthenocarpic expression in two tomato populations, which both contained *S. habrochaites* background. Inheritance of parthenocarpic expression in each population was controlled by two QTL with one QTL being common in both populations.

Due to the desirability of parthenocarpic fruits, the inheritance of parthenocarpic expression has been studied in a number of other species. A QTL study of parthenocarpic inheritance in eggplant revealed a two QTL model of inheritance (Miyatake et al., 2012). Analyses of segregation data found parthenocarpy to be under the control of at least two dominant genes in mandarin (Vardi et al., 2008). In pepino (*Solanum muricatum* Aiton), parthenocarpy was inherited as a single dominant gene (Prohens et al., 1998). Segregation ratios in diploid banana (*M. acuminata*) suggested the presence of at least three genes influencing the inheritance of parthenocarpy (Simmonds, 1953). Finally, observations of parthenocarpic expression in blueberry suggested parthenocarpy was complex and polygenically inherited (Ehlenfeldt and Vorsa, 2007).

In cucumber, the mode of genetic inheritance remains unresolved although highly successful greenhouse cultivars have been developed. Hawthorn and Wellington (1930) and Meshcherov and Juldasheva (1974) both reported models consisting of a single recessive gene for the inheritance of parthenocarpy. Pike and Peterson (1969) also developed a single gene model, although they reported parthenocarpy to be inherited as a single incompletely dominant gene. Kvasnikov et al. (1970) were the first to propose complex inheritance for parthenocarpy with a model consisting of many recessive genes. This was followed by a proposal by de Ponti and Garretson (1976) of an additive three gene inheritance model. Similarly, El-Shawaf and Baker (1981) found parthenocarpy to be quantitatively inherited with both additive and non-additive gene effects. Most recently, Sun et al. (2006) reported four major genomic regions associated with parthenocarpic expression with significant epistasis and large genotype×environment interactions.

Artificial parthenocarpy is induced through the application of exogenous phytohormones. Auxin, gibberellic acid (GA), cytokinin, and combinations of these are the most common phytohormones used to induce parthenocarpic expression (Gillaspy et al., 1993; Pandolfini, 2009; Vivian-Smith and Koltunow, 1999). A number of studies have also demonstrated the ability of auxin transport inhibitors to induce parthenocarpic expression (Beyer and Quebedeaux, 1974; Robinson et al., 1971; Serrani et al., 2010). In addition to these, the exogenous application of brassinosteroids was found to induce parthenocarpic fruit set in cucumber (Fu et al., 2008). Meanwhile, abscisic acid and ethylene are reported to have antagonistic roles in fruit set and parthenocarpic expression (Nitsch et al., 2009; Pascual et al., 2009; Vriezen et al., 2008; Wang et al., 2009a).

Auxin signaling is regulated by indole-3-acetic acid (IAA) and auxin response factor (ARF) transcription factors that act as inhibitors of auxin responsive genes (Leyser, 2006). ARF and IAA proteins form heterodimers that recognize auxin responsive genes (Goetz et al., 2007). It has been proposed that prior to pollination, IAA/ARF protein complexes repress fruit set genes (Goetz et al., 2006). Application or endogenous biosynthesis of auxin triggers the proteolytic degradation of IAA proteins, which results in the disintegration of the heterodimer complex and subsequently releases repression of auxin responsive genes (Dharmosiri and Estelle, 2004; Leyser, 2006; Woodward and Bartel, 2005). In addition, silencing or loss of function mutations to either ARF or IAA proteins removes their inhibitory effects, likely through failure to form the heterodimer complex, and induces parthenocarpic expression (Goetz et al., 2006; Goetz et al., 2007; de Jong et al., 2009; Wang et al., 2005). Interestingly, exogenous application of gibberellic acid induces parthenocarpic expression without altering the expression of auxin responsive genes (Vriezen et al., 2008). Further, exogenous application gibberellin biosynthetic inhibitors can block auxin induced parthenocarpic expression (Fuentes et al., 2012; Serrani et al., 2008; Serrani et al., 2010).

Further evidence that gibberellins, and not auxins, are critical to parthenocarpic fruit development comes from analysis of hormone levels between an obligatory parthenocarpic mandarin cultivar and a self-incompatible cultivar that exhibits stimulative parthenocarpy (Talon et al., 1990; Talon et al., 1992). Differences were observed in the levels of gibberellins throughout development between the cultivars (measurements taken 24 days before anthesis until 40 days after anthesis) with gibberellin concentrations reaching their highest levels at anthesis in the obligatory parthenocarpic cultivar. Gibberellin levels in the stimulated parthenocarpic cultivar were not changed from the control at anthesis. In addition, comparable levels of auxin

were observed between the two cultivars throughout the experiment. Similar reports from Arabidopsis and tomato consistently report elevated expression of gibberellins in association with parthenocarpic expression (Dorcey et al., 2009; Fos et al., 2000; Fos et al., 2001; Olimpieri et al., 2007; Pascual et al., 2009; Serrani et al., 2007; Serrani et al, 2008; Serrani et al., 2010). Ben-Cheikh et al. (1997) found that the increase in gibberellins observed during pollination was not the result of the pollen itself and suggests that another factor contained within pollen triggers gibberellin biosynthesis.

The gibberellin signaling pathway is regulated by inhibitory DELLA proteins that restrict plant growth and negatively regulate gibberellin growth responses (Dill and Sun, 2001; Dill et al., 2004; Li et al., 2012; Sun, 2011). DELLA proteins are members of the GRAS protein family of transcription factors characterized by the conserved amino acid motif "DELLA" (Thomas and Sun, 2004). The number of reported endogenous DELLA proteins varies by species with five identified in *Arabidopsis*, four identified in cotton, and only one identified in rice (Hu et al., 2011; Ueguchi-Tanaka et al., 2007). Silencing or loss of DELLA proteins has been found to induce facultative parthenocarpic expression (Carrera et al., 2012; Dorcey et al., 2009; Fuentes et al., 2012; Marti et al., 2007). Applications of auxin, cytokinin, and brassinosteroids have all been found to promote biosynthesis of gibberellins (Bouquin et al., 2001; Ding et al., 2013; Fuentes et al., 2012; Jager et al., 2005; Li et al., 2012; Nadhzimov et al., 1988; Serrani et al., 2008; Serrani et al., 2010; Wang et al., 2009b; Weiss and Ori, 2007). Increased gibberellin levels lead to degradation of DELLA proteins through binding of DELLAs in the GA-GID1-DELLA complex and releases DELLA mediated repression of GA responsive genes (Harberd et al., 2009, Sun, 2011). Interestingly, the ability of multiple hormones to affect gibberellin biosynthesis suggests a hierarchy in plant hormone signaling.

Parthenocarpic expression has been manipulated through the use of transgenes designed to overexpress auxin. This has been achieved through the expression of the *DefH9-iaaM* transgene construct in a variety of crops (Donzella et al., 2000; Ficcadenti et al., 1999; Mezzetti et al., 2004; Rotino et al., 1997; Yin et al., 2006). *DefH9* is a placenta-ovule specific promoter from *Antirrhinum majus* (Rotino et al., 1997). The *iaaM* gene of *Pseudomonas syringae* encodes tryptophan 2-monoxigenase, an enzyme converting tryptophan to indole-acetamide, which is spontaneously or enzymatically converted to indole-3-acetic-acid (auxin) within plant cells (Kosuge et al., 1966; Pandolfini et al., 2009; Rotino et al., 1997). *DefH9-iaaM* containing plants exhibit facultative parthenocarpy and seed set is possible (Rotino et al, 2005). However, ovary development commences prior to anthesis in *DefH9-iaaM* plants and most fruits develop parthenocarpically (Acciarri et al., 2002; Rotino et al., 2005). A second transgenic construct consisting of the ovary and young fruit specific *TPRP-F1* promoter and the *Agrobacterium rhizogenes* gene *rolB* has been used. The *rolB* gene conditions increased sensitivity to auxin (Carmi et al., 2003). The transgenic fruits created with both constructs have been reported to have equal or improved quality when compared to their seeded counterparts (Carmi et al., 2003; Costantini et al., 2007; Maestrelli et al., 2003; Rotino et al., 2005).

Parthenocarpic cucumber cultivars have been successfully developed, but the genetic and molecular mechanisms behind parthenocarpic expression remain largely unknown. This information is essential for breeding programs proposing to incorporate parthenocarpy into elite processing cucumber populations and hybrids. Therefore, this study was designed to better define what is considered true parthenocarpic expression and then to use this knowledge to identify QTL associated with parthenocarpic fruit set. A new approach to phenotyping parthenocarpic fruit set in cucumber was implemented in order to better define true

parthenocarpic expression. This new approach sought to build off of studies demonstrating that parthenocarpic fruit set is determined in the days before and immediately after anthesis by focusing on early fruit initiation and development. With a clear approach to phenotypic evaluation, traditional QTL mapping approaches such as interval mapping, composite interval mapping, and multiple interval mapping were employed to detect and construct optimal models for the inheritance of parthenocarpic fruit set in cucumber. With the identification of genomic regions known to associate with parthenocarpic fruit set, bioinformatic analyses of these regions were conducted and potential candidate genes for parthenocarpic fruit set were identified. The QTL identified for parthenocarpic fruit set by this study are valuable to cucumber breeders interested in developing parthenocarpic cultivars and to researchers interested in the genetic and molecular mechanisms of parthenocarpic fruit set.

Literature Cited

Acciarri, N., F. Restaino, G. Vitelli, D. Perrone, M. Zottini, T. Pandolfini, A. Spena, and G.L. Rotino. 2002. Genetically modified parthenocarpic eggplants: improved fruit productivity under both greenhouse and open field cultivation. BMC Biotechnol. 2:4.

Ben-Cheikh, W., J. Perz-Botella, F.R. Tadeo, M. Talon, E. Primo-Millo. 1997. Pollination increases gibberellin levels in developing ovaries of seeded varieties of *Citrus*. Plant Physiol. 114:557-564.

Beyer, E.M. Jr. and B. Quebedeaux. 1974. Parthenocarpy in cucumber: mechanism of action of auxin transport inhibitors. J. Amer. Soc. Hort. Sci. 99:385-390.

Bouquin, T., C. Meier, R. Foster, M.E. Nielsen, and J. Mundy. 2001. Control of specific gene expression by gibberellin and brassinosteroid. Plant Physiol. 127:450-458.

Carbonell, J. and J.L. Garcia-Martinez. 1980. Fruit-set of unpollinated ovaries of *Pisium sativum* L.: influence of vegetative parts. Planta. 147:444-450.

Carmi, N., Y. Salts, B. Dedicova, S. Shabtai, and R. Barg. 2003. Induction of parthenocarpy in

tomato via specific expression of the *rolB* gene in the ovary. Planta. 217:726-735.

Carrera, E., O. Ruiz-Rivero, L.E.P. Peres, A. Atares, and J.L. Garcia-Martinez. 2012. Characterization of the *procera* tomato mutant shows novel functions of the SIDELLA protein in the control of flower morphology, cell division and expansion, and the auxin-signaling pathway during fruit-set and development. Plant Physiol. 160:1581-1596.

Coombe, B.G. 1962. The effect of removing leaves, flowers and shoot tips on fruit-set in *Vitis vinifera* L. J. Hortic. Sci. 37:1-15.

Coombe, B.G. 1976. The development of fleshy fruits. Annu. Rev. Plant Physiol. 27:507-528.

Costantini, E., L. Landi, O. Silvestroni, T. Pandolfini, A. Spena, and B. Mezzetti. 2007. Auxin synthesis-encoding transgene enhances grape fecundity. Plant Physiol. 143:1689-1694.

Dharmosiri, N. and M Estelle. 2004. Auxin signaling and regulated protein degradation. Trends Plant Sci. 9:302-308.

Dill, A. and T.P. Sun. 2001. Synergistic derepression of gibberellin signaling by removing RGA and GAI function in *Arabidopsis thaliana*. Genetics. 159:777-785.

Dill, A., S.G. Thomas, J. Hu, C.M. Steber, and T.P. Sun. 2004. The *Arabidopsis* F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. Plant Cell. 16:1392-1405.

Ding, J., B. Chen, X. Xia, W. Mao, K. Shi, Y. Zhou, and J. Yu. 2013. Cytokinin-induced parthenocarpic fruit development in tomato is partly dependent on enhanced gibberellin and auxin biosynthesis. PLoS One. 8:e70080.

Donzella, G., A. Spena, and G.L. Rotino. 2000. Transgenic parthenocarpic eggplants: superior germplasm for increased winter production. Mol. Breeding. 6:79-86.

Dorcey, E., C. Urbez, M.A. Blazquez, J. Carbonell, and M.A. Perez-Amador. 2009. Fertilization-dependent auxin response in ovules triggers fruit development through the modulation of gibberellin metabolism in *Arabidopsis*. Plant J. 58:318-332.

Ehlenfeldt, M.K. and N. Vorsa. 2007. Inheritance patterns of parthenocarpic fruit development in highbush blueberry (*Vaccinium corymbosum* L.). HortScience. 42:1127-1130.

El-Shawaf, I.I.S. and L.R. Baker. 1981. Inheritance of parthenocarpic yield in gynoecious pickling cucumber for once-over mechanical harvest by diallel analysis of six gynoecious lines. J. Amer. Soc. Hort. Sci. 106:365-370.

Falavinga, A., M. Badino, and G.P. Soressi. 1978. Potential of the monomendelian factor *pat* in the tomato breeding for industry. Genetica Agraria. 32:160.

Ficcadenti, N., S. Sestili, T. Pandolfini, T. Crillo, G.L. Rotino, and A. Spena. 1999. Genetic engineering of parthenocarpic fruit development in tomato. Mol. Breeding. 5:463-470.

Fitting, H. 1909. Die Beeinflussung der Orchideenbluten durch die Bestaubung und durch andere Umstande. Zeitschrift fuer Botanik. 1:1-86.

Fos, M., F. Nuez, and J.L. Garcıa-Martınez. 2000. The gene pat-2, which induces natural parthenocarpy, alters the gibberellin content in unpollinated tomato ovaries. Plant Physiol. 122:471-480.

Fos, M., K. Proano, F. Nuez, and J.L. Garcia-Martinez. 2001. Role of gibberellins in parthenocarpic fruit development induced by the genetic system *pat-3*/*pat-4* in tomato. Physiol. Plantarum. 111:545-550.

Fu, F.Q., W.H. Mao, Y.H. Zhou, T. Asami, and J.Q. Yu. 2008. A role of brassinosteroids in early fruit development in cucumber. J. Exp. Bot. 59:2299-2308.

Fuentes, S., K. Ljung, K. Sorefan, E. Alvey, N.P. Harberd, and L. Ostergaard. 2012. Fruit growth in *Arabidopsis* occurs via DELLA-dependent and DELLA-independent gibberellin responses. Plant Cell. 24:3982-3996.

George, W., J. Scott, and W. Spilttstoesser. 1984. Parthenocarpy in tomato. Horticultural Reviews. 6:65-84.

Gillaspy, G., H. Ben-David, and W. Gruissem. 1993. Fruits: A developmental perspective. Plant Cell. 5:1439-1451.

Goetz, M., A. Vivian-Smith, S.D. Johnson, and A.M. Koltunow. 2006. AUXIN RESPONSE FACTOR 8 is a negative regulator of fruit initiation in Arabidopsis. Plant Cell. 18:1873-1886.

Goetz, M., L.C. Hooper, S.D. Johnson, J.C. Rodrigues, A. Vivian-Smith, and A.M. Koltunow. 2007. Expression of aberrant forms of AUXIN RESPONSE FACTOR 8 stimulates parthenocarpy in *Arabidopsis* and tomato. Plant Physiol. 145:351-366.

Gorguet, B., A.W. van Heusden, and P. Lindhout. 2005. Parthenocarpic fruit development in tomato. Plant Biol. 7:131-139.

Gorguet, B., P.M. Eggink, J. Ocana, A. Tiwari, D. Schipper, R. Finkers, R.G.F. Visser, and A.W. van Heusden. 2008. Mapping and characterization of novel parthenocarpy QTLs in tomato. Theor. Appl. Genet. 116:755-767.

Gustafson, F.G. 1937. Parthenocarpy induced by pollen extracts. Am. J. Bot. 24:102-107.

Gustafson, F.G. 1942. Parthenocarpy: natural and artificial. Bot. Rev. 8:599-654.

Harberd, N.P., E. Belfield, and Y. Yasumura. 2009. The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: how and "inhibitor of an inhibitor" enables flexible response to fluctuating environments. Plant Cell. 21:1328-1339.

Hawthorn, L.R. and R. Wellington. 1930. Geneva, a greenhouse cucumber that develops fruit without pollination. NY (Geneva) Agric Exp. Stat. Bull. 580:223-234.

Hu, M.Y., M. Luo, Y.H. Xiao, X.B. Li, K.L. Tan, L. Hou, J. Dong, D.M. Li, S.Q. Song, J. Zhao, Z.L. Zang, B.L. Li, and Y. Pei. 2011. Brassinosteroids and auxin down-regulate DELLA genes in fiber initiation and elongation of cotton. Agric. Sci. China. 10:1168-1176.

Jager, C.E., G.M. Symons, J.J. Ross, J.J. Smith, and J.B. Reid. 2005. The brassinosteroid growth response in pea is not mediated by changes in gibberellin content. Planta. 221:141-148.

de Jong, M., M. Wolters-Arts, J.L. Carcia-Martinez, C. Mariani, and W.H. Vriezen. 2009. The *Solanum lycopersicum* auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development Plant J. 57:160-170.

Kvasnikov, B.V., N.T. Rogova, S.I. Tarakonova, and I. Ignatova. 1970. Methods of breeding vegetable crops under the covered ground. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 42:45-57.

Kosuge, T., M.G. Heskett, and E.E. Wilson. 1966. Microbial synthesis and degradation of the indole-3-acetic-acid. J. Biol. Chem. 241:3738-3744.

Leyser, O. 2006. Dynamic integration of auxin transport and signaling. Curr. Biol. 16:424-433.

Li, Q.F., C. Wang, L. Jiang, S. Li, S.S.M. Sun, and J.X. He. 2012. An interaction between BZR1 and DELLAs mediates direct signaling crosstalk between brassinosteroids and gibberellins in *Arabidopsis*. Sci. Signal. 5:ra72.

Maestelli, A., R. Lo Scalzo, G.L. Rotino, N. Acciarri, A. Spena, G. Vitelli, and G. Bertolo. 2003. Freezing effect on some quality parameters of transgenic parthenocarpic eggplants. J. Food Eng. 56:285-287.

Marti, C., D. Orzaez, P. Ellul, V. Moreno, J. Carbonell, and A. Granell. 2007. Silencing of *DELLA* induces facultative parthenocarpy in tomato fruits. Plant J. 52:865-876.

Meshcherov, E.T. and L.W. Juldasheva. 1974. Parthenocarpy in cucumber. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 51:204-213.

Mezzetti, B., T. Pandolfini, G.L. Rotino, V. Dani, and A. Spena. 2004. The *DefH9-iaaM* auxin-synthesizing gene increases plant fecundity and fruit production in strawberry and raspberry. BMC Biotechnol. 4:4.

Miyatake, K., T. Saito, S. Negoro, H. Yamaguchi, T. Nunome, A. Ohyama, and H. Fukuoka.

2012. Development of selective markers linked to a major QTL for parthenocarpy in eggplant (*Solanum melongena* L.). Theor. Appl. Genet. 124:1403-1413.

Nadhzimov, U.K., S.C. Jupe, M.G. Jones, and I.M. Scott. 1988. Growth and gibberellin relations of the extreme dwarf d^x tomato mutant. Physiol. Plant. 73:252-256.

Nitsch, L.M., C. Oplaat, R. Feron, Q. Ma, M. Wolters-Arts, P. Hedden, C. Mariani, and W.H. Vriezen. 2009. Abscisic acid levels in tomato ovaries are regulated by *LeNCED1* and *SlCYP707A1*. Planta. 229:1335-1346.

Noll, F. 1902. Fruchtbilding ohne vorausgegangene Bestaubung (Pathenocarpie) bei der Gurke. Gesellschaft fur Natur-u. Heilkunde zu Bonn. 106:149-162.

Nuez, F., J. Costa, and J. Cuartero. 1986. Genetics of the parthenocarpy for tomato varieties "Sub-Artic Plenty", "75/59", and "Severianin". Zeitschrift fur Pflanzenzuchtung. 96:200-206.

Olimpieri, I., F. Silicato, R. Caccia, L. Mariotti, N. Ceccarelli, G.P. Soressi, and A. Mazzucato. 2007. Tomato fruit set driven by pollination or by the parthenocarpic fruit allele are mediated by transcriptionally regulated gibberellin biosynthesis. Planta. 226:877-888.

Pandolfini, T. 2009. Seedless fruit production by hormonal regulation of fruit set. Nutrients. 1:168-177.

Pandolfini, T., B. Molesini, and A. Spena. 2009. Parthenocarpy in crop plants. *In* L. Ostergaard, ed, Fruit development and seed dispersal. Wiley Blackwell. Oxford, United Kingdom.

Parry, M.S. 1976. Effect of pruning and partial disbudding on the cropping of Cornice and Beurre Hardy pears. J. Hortic. Sci. 51:159-166.

Pascual. L., J.M. Blanca, J. Canizares, and F. Nuez. 2009. Transcriptomic analysis of tomato carpel development reveals alterations in ethylene and gibberellins synthesis during *pat3/pat4* parthenocarpic fruit set. BMC Plant Biol. 9:67.

Picken, A.J.F. 1984. A review of pollination and fruit set in the tomato (*Lycopersicon esculentum* Mill.) J. Hort. Sci. 59:1-13.

Pike, L.M. and C.E. Peterson. 1969. Inheritance of parthenocarpy in the cucumber (*Cucumis sativus* L.) Euphytica. 18:101-105.

de Ponti, O.M.B. and F. Garretsen. 1976. Inheritance of parthenocarpy in pickling cucumbers (*Cucumis sativus* L.) and linkage with other characters. Euphytica. 25:633-642.

Philouze, J. and B. Maisonneuve. 1978. Heredity of the natural ability to set parthenocarpic fruits in the Soviet variety 'Severianin'. Tomato Genet. Coop. Rep. 28:12-13.

Philouze, J. and P. Pecaut. 1986. Natural parthenocarpy in tomato. III. Study of the parthenocarpy due to the gene *pat* (*parthenocarpic fruit*) in the line Montfavet 191. Agronomie. 6:243-248.

Philouze, J., M. Buret, F. Duprat, M. Nicolas-Grotte, and J. Nicolas. 1988. Caracteristiques agronomiques et physic-chimiques de lignees de tomates isogeniques, sauf pour le gene *pat-2* de parthenocarpie, dans trios types varietaux, en culture de printemps, sous serre plastique tres peu chauffee. Agronomie. 8:817-823.

Philouze, J. 1989. Natural parthenocarpy in tomato. IV. A study of the polygenic control of parthenocarpy in line 75/59. Agronomie. 9:63-75.

Prohens, J., J.J. Ruiz, and F. Nuez. 1998. The inheritance of parthenocarpy and associated traits in pepino. J. Am. Soc. Hort. Sci. 123:376-380.

Quinlan, J.D. and A.P. Preston. 1971. The influence of shoot competition on fruit retention and cropping of apple trees. J. Hortic. Sci. 46:525-534.

Rotino, G.L., E. Perri, M. Zottini, H. Sommer, and A. Spena. 1997. Genetic engineering of parthenocarpic plants. Nat. Biotechnol. 15:1398-1401.

Rotino, G.L., N. Acciarri, E. Sabatini, G. Mennella, R. Lo Scalzo, A. Maestrelli, B. Molesini, T. Pandolfini, J. Scalzo, B. Mezzetti, and A. Spena. 2005. Open field trial of genetically modified parthenocarpic tomato: seedlessnes and fruit quality. BMC Biotechnol. 5:32.

Saunders, R.C., G. Jacobs, and D.K. Strydom. 1991. Effect of pruning on fruitset and shoot growth of ‘Packham’s Triumph’ pear trees. Sci. Hort. 47:239-245.

Schwabe, W.W. and J.J. Mills. 1981. Hormones and parthenocarpic fruit set: a literature survey. Horticultural Abstracts. 51:661-699.

Serrani, J.C., M. Fos, A. Atares, and J.L. Garcia-Martinez. 2007. Gibberellin regulation of fruit-set and growth in tomato. Plant Physiol. 145:246-257.

Serrani, J.C., O. Ruiz-Rivero, M. Fos, and J.L. Garcia-Martinez. 2008. Auxin-induced fruit-set in tomato is mediated in part by gibberellins. Plant J. 56:922-934.

Serrani, J.C., E. Carrera, O. Ruiz-Rivero, L. Gallego-Giraldo, L.E.P. Peres, and J.L. Garcia-Martinez. 2010. Inhibition of auxin transport from the ovary or from the apical shoot induces parthenocarpic fruit-set in tomato mediated by gibberellins. Plant Physiol. 153:851-862.

Simmonds, N.W. 1953. Segregations in some diploid bananas. J. Genet. 51:458-469.

Soressi, G.P. and F. Salamini. 1975. A monomendelian gene inducing parthenocarpic fruits. Tomato Genet. Coop. Rep. 25:22.

Srivastava, L.M. 2002. Plant Growth and Development: Hormones and Environment. Elsevier Science. San Diego, CA.

Sun, Z., J.E. Staub, S.M. Chung, and R.L. Lower. 2006. Identification and comparative analysis of quantitative trait loci (QTL) associated with parthenocarpy in processing cucumber. Plant Breeding. 125:281-287.

Sun, T.P. 2011. The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants. Curr. Biol. 21:R338-R345.

Talon, M., L. Zacarias, and E. Primo-Millo. 1990. Hormonal changes associated with fruit set and development in mandarins differing in their parthenocarpic ability. Plant Physiol. 79:400-406.

Talon, M., L. Zacarias, and E. Primo-Millo. 1992. Gibberellins and parthenocarpic ability in developing ovaries of seedless mandarins. Plant Physiol. 99:1575-1581.

Tsao, T. 1980. Growth substances: role in fertilization and sex expression. Plant Growth Substances. Springer-Verlag. New York, NY.

Ueguchi-Tanaka, M. Nakajima, E. Katoh, H. Ohmiya, K. Asano, S. Saji, X. Hongyu, M. Ashikari, H. Kitano, I. Yamaguchi, and M. Matsuoka. 2007. Molecular interactions of a soluble gibberellin receptor GID1, with a rice DELLA protein, SLR1, and gibberellin. Plant Cell. 19:2140-2155.

Vardi, A., I. Levin, and N. Carmi. 2008. Induction of seedlessness in citrus: from classical techniques to emerging biotechnological approaches. J. Am. Soc. Hortic. Sci. 133:117-126.

Vardy, E., D. Lapushner, A. Genizi, and J. Hewitt. 1989a. Genetics of parthenocarpy in tomato under a low temperature regime I. Euphytica 41:1-8.

Vardy, E., D. Lapushner, A. Genizi, and J. Hewitt. 1989b. Genetics of parthenocarpy in tomato under a low temperature regime II. Cultivar ‘Severianin’. Euphytica 41:9-15.

Vivian-Smith, A. and A.M. Koltunow. 1999. Genetic analysis of growth regulator-induced parthenocarpy in *Arabidopsis*. Plant Physiol. 121:437-451.

Vriezen, W.H., R. Feron, F. Maretto, J. Keijman, and C. Mariani. 2008. Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. New Phytol. 177:60-76.

Wang, H., B. Jones, Z. Li, P. Frasse, C. Delalande, F. Regad, S. Chaabouni, A. Latche, J.C. Pech, and M. Bouzayen. 2005. The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. Plant Cell. 17:2676-2692.

Wang, H., N. Schauer, B. Usadel, P. Frasse, M. Zouine, M. Hernould, A. Latche, J.C. Pech, A.R. Fernie, and M. Bouzayen. 2009a. Regulatory features underlying pollination-dependent and –independent tomato fruit set revealed by transcript and primary metabolite profiling. Plant Cell. 21:1428-1452.

Wang, L, Z. Wang, Y. Xu, S.H. Joo, S.K. Kim, Z. Xue, Z. Xu, Z. Wang, and K. Chong. 2009b. *OsGSR1* is involved in crosstalk between gibberellins and brassinosteroids in rice. Plant J. 57:498-510.

Weiss, D. and N. Ori. 2007. Mechanisms of cross talk between gibberellin and other hormones. Plant Physiol. 144:1240-1246.

Westwood, M.N. and H.O. Bjornstad. 1974. Fruitset as related to girdling, early cluster thinning and pruning of 'Anjou' and 'Comice' pear. HortScience. 9:342-344.

Woodward, A.W. and B. Bartel. 2005. Auxin: regulation, action and interaction. Ann. Bot-London. 95:707-735.

Yasuda, S. 1930. Parthenocarpy induced by the stimulus of pollination in some plants of the Solanaceae. Agriculture and Horticulture. 5:287-294.

Yasuda, S. 1934. Parthenocarpy caused by the stimulus of pollination in some plants of Solanaceae. Agriculture and Horticulture. 9:647-656.

Yasuda, S. 1935. Parthenocarpy induced by the stimulation of pollination in some plants of the Cucurbitaceae. Agriculture and Horticulture. 10:1385-1390.

Yin, Z., R. Malinowski, A. Ziolkowska, H. Sommer, W. Plader, and S. Malepszy. 2006. The *DefH9-iaaM*-containing construct efficiently induces parthenocarpy in cucumber. Cell. Mol. Biol. Lett. 11:279-290.

Chapter 1

A Novel Approach to Phenotypic Evaluation of Parthenocarpic Fruit Set in Processing Cucumber (*Cucumis sativus* L.)

Abstract

Parthenocarpic processing cucumber (*Cucumis sativus* L.) varieties have the potential for increasing yield, improving fruit quality, and extending production periods. Since parthenocarpy is often considered a yield component, it is difficult to separate the true parthenocarpic character from other yield related traits. In order to better define what is considered true parthenocarpic expression, a new approach to phenotyping parthenocarpic fruit set in cucumber was implemented focusing on early fruit initiation and development. An $F_{2:3}$ population was used to characterize the inheritance of parthenocarpic fruit set by crossing a highly parthenocarpic inbred line, '2A', with a non-parthenocarpic inbred line, 'Gy8'. A continuous distribution of F_3 family means suggested that parthenocarpic fruit set is quantitatively inherited in this population. Patterns of fruit set on experimental plants revealed that potential for parthenocarpic fruit set could be effectively evaluated with as few as five pistillate flowers. In addition, the pruning of axillary shoots in the maintenance of greenhouse plants was found to inadvertently increase parthenocarpic fruit set.

Introduction

Cucumber (*Cucumis sativus* L.) is one of the most cultivated and economically important crops in the world. Worldwide production in 2010 was estimated at 62.7 million metric tons while the US crop totaled nearly 900,000 metric tons with a farm gate value of $462 million (FAOSTAT, 2013). Despite the crops prevalence, processing cucumber yields in the U.S. have not substantially increased from levels seen in the 1980's (Gusmini and Wehner, 2008). A phenomenon known as first fruit inhibition, where the first fertilized fruit inhibits growth of subsequent fruits, is thought to be a major obstacle to yield improvement in cucumber (Denna, 1973; El-Shawaf and Baker, 1981; McCollum, 1934; de Ponti, 1976; Strong, 1921; Sun et al., 2006a; Tiedjens, 1928; Uzcategui and Baker, 1979). A potential solution is the use of gynoecious (pistillate flowers only) parthenocarpic cucumber varieties. The use of gynoecy is essential for successful parthenocarpic cultivars since genetic parthenocarpy in cucumber is facultative and plants that are pollinated are able to set seeded fruit (Denna, 1973). Gynoecious parthenocarpic varieties offer several advantages over conventional seeded varieties. One beneficial factor is that parthenocarpic varieties are able to set fruit sequentially without suffering from first fruit inhibition (Denna, 1973; Sun et al., 2006a). Second, parthenocarpic varieties do not require pollination and are therefore less vulnerable to poor pollination conditions (abiotic and biotic) and the need for insect pollinators (de Ponti, 1976; Pike and Peterson, 1969; Sun et al., 2006a; Varoquaux et al., 2000). Third, parthenocarpic varieties often have more uniformly shaped fruit desired by the processing industry because they do not suffer from incomplete pollination that can cause misshapen fruit in conventional varieties (Aalbersberg and van Wijchen, 1987; Baker et al., 1973; de Ponti, 1976).

Parthenocarpy, defined as the development of virgin fruits, is a desired trait in many plant species. Parthenocarpy has long been an important trait in cucumber, especially for greenhouse production (Sturtevant, 1890). European greenhouse cultivars in the 19th century were selected for high yield, but often cultivars were also indirectly selected for parthenocarpic fruit set due to their increased productivity in poor pollinating conditions (Robinson and Reiners, 1999). Since then, parthenocarpic fruit set in cucumber has been manipulated both genetically and with the exogenous application of various synthetic phytohormones. The application of auxin, cytokinin, gibberellic acid, brassinosteroids, and auxin transport inhibitors all result in the induction of parthenocarpic fruit set in cucumber (Beyer and Quebedeaux, 1974; Cantliffe and Phatak, 1975; Choudhury and Phatak, 1959; Elassar et al., 1974; Fu et al., 2008; Homan, 1964; Kim et al., 1992; Robinson et al., 1971). However, this method of obtaining parthenocarpic expression has many drawbacks including the need for continuous application of phytohormones throughout growth, increased input costs for growers, environmental impact concerns, and human dietary concerns related to consumption of phytohormones (Rotino et al., 1997).

Since parthenocarpy can be easily induced by the application of various hormones, many have suggested the mechanisms for parthenocarpic fruit set are hormone production, transport, and/or crosstalk related. Genetic studies have been inconclusive on the inheritance of parthenocarpy in cucumber and have ranged from proposals of single gene inheritance to complex multigenic inheritance models (El-Shawaf and Baker, 1981; Hawthorn and Wellington, 1930; Kvasnikov et al., 1970; Meshcherov and Juldasheva, 1974; Pike and Peterson, 1969; de Ponti and Garretson, 1976; Sun et al., 2006a; Sun et al., 2006b). A recent study by Sun et al. (2006b) concludes with a model of complex genetic inheritance that is heavily influenced by

environmental conditions. Parthenocarpy has also been induced with the use of transgenic overexpression of auxin in cucumber ovaries (Yin et al., 2006).

The objective of this research is to gain a better understanding of the genetic characteristics of parthenocarpy in cucumber so that it may be better utilized by breeding programs seeking to develop parthenocarpic processing cucumber cultivars. Since parthenocarpy is often considered a yield component, it is difficult to separate the true parthenocarpic character from other yield related traits. A new approach to phenotyping parthenocarpic fruit set in cucumber was implemented in order to better define true parthenocarpic expression. This new approach sought to build off studies demonstrating that parthenocarpic fruit set is determined in the days before and immediately after anthesis by focusing on early fruit initiation and development (Fos et al., 2000, Gillaspy et al., 1993; Molesini et al., 2009; Pascual et al., 2009; Ruan et al., 2012).

Materials and Methods

Population Development

Two gynoecious U.S. processing cucumber inbred lines differing in their expression of parthenocarpic fruit set were selected for the study of parthenocarpic fruit set (Addendum 1). The highly parthenocarpic inbred line, '2A', is gynoecious (gy), indeterminate (De), and is able to consistently set multiple parthenocarpic fruits without pollination in both open field and greenhouse environments (Sun et al., 2006a). The non-parthenocarpic inbred line, 'Gy8', is gynoecious, indeterminate, and yields few to no fruit in the absence of pollination (Sun et al., 2006a). 'Gy8' was selected because it exhibits growth and fruit characteristics similar to '2A'

including: stable gynoecious expression, fresh and brine stock quality, an indeterminate growth habit, and a blocky shape with length to diameter ratios greater than three when fruits are four cm in diameter (Sun, 2004). An additional benefit to the selection of these parental lines was the opportunity to directly compare results to previous research on parthenocarpic yield in cucumber performed with another 2A×Gy8 $F_{2:3}$ population developed by Sun et al. (2006a and 2006b).

An $F_{2:3}$ population was developed to explore the complex inheritance of parthenocarpic fruit set in cucumber. Inbred lines '2A' and 'Gy8' were crossed to produce F_1 seed. A single F_1 plant was self-pollinated and 205 F_2 plants were grown in a greenhouse during the spring of 2011. Each F_2 plant was self-pollinated to produce F_3 seeds which were used to generate the F_3 families described in this study. Since all plants mentioned in this study are gynoecious; parental, F_1, and F_2 plants involved in the creation of the population by seed were manipulated to produce male flowers for use in pollination with 3mM silver thiosulfate $[Ag(S_2O_3)_2]^{3-}$ as a foliar spray (Nijs and Visser, 1980).

Experimental Organization

Experiment 1

Experiment 1 was conducted from July 2011 to September 2011 at the University of Wisconsin-Madison Walnut Street Research Greenhouses (WSGH) located in Madison, WI. Five greenhouses, each measuring 6.1 m × 6.1 m were used in order to accommodate the large number of plants. Each greenhouse contained one plant from each of 201 F_3 families, four plants from each of the parental inbred lines '2A' and 'Gy8' and one F_1 plant. Each greenhouse contained 210 plants. In total, experiment 1 included 1050 plants allocated as 5 plants from each of 201 F_3 families, 20 plants from each of '2A' and 'Gy8', and 5 F_1 plants.

Plants in each greenhouse were placed in staggered rows oriented in a north to south direction with 14 plants in each row and a total of 15 rows. The diameter of each potted plant was 25.4 cm with 12.7 cm of space between pots within individual rows and 14.2 cm of space between pots of neighboring rows. To facilitate access to the plants for watering and care, two 45.7 cm walkways were created between the 5^{th} and 6^{th} rows and the 10^{th} and 11^{th} rows, respectively. The space between each of the walls of the greenhouse and the plants was 45.7 cm (Addendum 2).

Experiments 2 and 3

To address issues related to crowding found in experiment 1, some modifications were made in experiments 2 and 3. Experiment 2 was conducted from September 2011 to December 2011 and experiment 3 was conducted from March 2012 to June 2012. Both experiments were conducted in five 6.1 m × 6.1 m greenhouses at WSGH. Three plants from each of 205 F_3 families were randomly distributed across each experiment with an additional four plants from each of '2A' and 'Gy8' and one F_1 plant included in each greenhouse as controls. Each greenhouse contained 132 plants. In total, experiment 2 and 3 each included 660 plants allocated as 3 plants from each of 205 F_3 families, 20 plants from each of '2A' and 'Gy8', and 5 F_1 plants.

Plants in each greenhouse were placed in rows oriented in a north to south direction with 11 plants in each row and a total of 12 rows. The diameter of each potted plant was 25.4 cm with 22.9 cm of space between pots within individual rows and 17.8 cm of space between pots of neighboring rows. To facilitate access to the plants for watering and care, two 45.7 cm walkways were created between the 4^{th} and 5^{th} rows and the 8^{th} and 9^{th} rows respectively. The space between each of the walls of the greenhouse and the plants was 45.7 cm (Addendum 3).

Greenhouse Conditions and Plant Maintenance

Plants were grown in Classic 1000 plastic pots (20.6 cm bottom x 25 cm top x 23.2 cm height with a 9.5 liter capacity) with Metro Mix Professional Growing Mix soil (Sun Gro Horticulture Canada CM Ltd.). Along with the starter nutrients included in the Metro Mix Professional Growing Mix, plants were supplemented with 70.9 g of Nutricote 100 Controlled Release Fertilizer (Arysta LifeScience North America, Cary, NC). Each greenhouse was environmentally controlled throughout the duration of each experiment. Plants were grown under 14 hour days. When natural light levels dropped below 650 µE, supplemental artificial high pressure sodium lights were utilized. Temperatures were maintained at 29.4°C during day time hours and 23.9°C during night time hours. During growth, plants were watered once per day and regularly staked to grow vertically on 1.83 m long bamboo poles with wire twist ties.

Data Collection

Parthenocarpic fruit set was measured as the number of fruits initiated on each plant. Ovaries were considered to have initiated development if upon visual inspection, clear growth and expansion was visible. Ovaries that had initiated growth but later ceased at any point during development were included as successfully initiated fruits (Figure 1). In order to limit confounding factors related to other traits, plants were maintained as follows:

1. When plants were between 10-15 nodes in length, the first five nodes of each plant were cleared of all vegetation, including flowers, to aid in limiting differences in flowering time (development of flowers at earlier nodes) and problems in subsequent fruit setting associated

with crown fruit set (Denna, 1973). This also served to create space for watering and air circulation.

2. When plants were between 10-15 nodes in length, all lateral branches were removed and continued to be removed regularly throughout the remainder of growth in an attempt to equalize potential differences in plant photosynthetic capacity.

3. Plants were inspected regularly to only allow one pistillate flower per node to ensure that each plant had an equal number of pistillate flowers and opportunity for fruit set.

The presence of an initiated fruit and the size of the fruit if present were recorded from individual plant nodes 6 through 30. Data collection was conducted when approximately 95% of plants had reached 35 nodes in length (approximately 60 days after germination). When plants had reached 35 nodes in length, conclusive evaluation of the 30^{th} node could be made in data collection.

Data Analysis

Mean, median, maximum, and minimum values for initiated fruit in each experiment were calculated from data collected for all F_3 individuals. Frequency distributions were calculated from all F_3 individuals in each experiment and also from F_3 family means for comparison. Due to the lack of replication and sampling errors related to the growth of only three to five segregating F_3 individuals from each F_3 family, an examination of genetic and environmental effects was not performed. To determine if greenhouse environments were similar between experiments, two sample t-tests were calculated between pairs of experiments from experiment means calculated with data from all F_3 individuals. In addition, heat maps were constructed for

each experiment in order to visually inspect for major spatial patterns suggesting the uneven distribution of values. A Spearman rank correlation was performed in order to determine if data collected in each individual experiment could be pooled in order to alleviate the severity of sampling errors from only sampling three to five individuals of each F_3 family. Data analysis and plots were created with the statistical software R version 2.13.2.

Supplemental Validation Experiment: Effects of Plant Maintenance and Treatment

An experiment observing the effects of plant treatment in experiments 1-3 was grown at WSGH consisting of 20 total plants with 10 plants coming from each of the parental inbred lines '2A' and 'Gy8'. Five plants from each parental line were subjected to the same plant maintenance and greenhouse conditions as used in the focus study. The remaining five plants from each parental line were allowed to grow unhindered in the same greenhouse environment. Parthenocarpic fruit set data was recorded when all plants had reached 35 nodes in main stem length. For plants that had been subjected to plant maintenance, data was collected only from nodes 6 through 30 on the main stem. For the plants allowed to grow unhindered, data was collected first from nodes 1 through 30 of the main stem with inclusion of all lateral branches. A second data collection was taken from only nodes 6 through 30 of each plant.

Results and Discussion

General Assessment of the Three Focus Experiments

The three focus experiments of this study were assessed for overall data quality. Each of the three greenhouse experiments conducted in this study returned comparable ranges for the number

of fruit initiated per plant (0-15) and average number of fruit initiated per plant (3.23-3.76) (Table 1). However, two sample t-tests conducted between pairs of experiments revealed that while the average number of fruit initiated in experiments 2 and 3 are not significantly different, they both varied significantly from experiment 1 (Table 2). In addition, the frequency distribution of total parthenocarpic fruit initiated in experiment 1 showed some skewing of data towards fewer fruit when compared with experiments 2 and 3 (Figure 2).

These results complement observations made during the growth of the experimental plants. During experiment 1, observations were made that some plants were losing foliage and failing to set fruit on lower plant nodes. The plants appeared to be suffering from crowding; presumed to be due to the high density planting used in experiment 1 (Addendum 2). This was unexpected based upon greenhouse observations made prior to this study. However, experiment 1 was completed as originally designed since many plants had already aborted flowers on the lower plant nodes and the severity of crowding was not deemed critical enough to compromise the experiment. Upon completion of experiment 1, a heat map did not reveal any major spatial patterns suggesting the uneven distribution of values (Addendum 4).

Before beginning experiment 2, the experimental design was modified to reduce the number of plants in the subsequent experiments by 40% to ensure that crowding would no longer be an interfering variable (Addendum 3). The number of F_3 families included in the study was increased from 201 to 205 simply to balance each greenhouse with even rows. No symptoms of plant crowding were observed during the growth of experiments 2 and 3. Both experiments 2 and 3 returned data that was highly consistent and a t-test concluded that the means of the two experiments are not significantly different (Table 1, Table 2). Heat maps of both experiments do

not reveal any major spatial patterns suggesting the uneven distribution of values (Addendum 5, Addendum 6).

Given that the two sample t-tests found experiment 1 to be significantly different from the other experiments, a Spearman rank correlation between each of the experiments was performed in pairs (Table 3). The Spearman rank correlations were performed using the F_3 family means obtained in each experiment. However, since F_3 plants are known to be genetically segregating, the data is susceptible to sampling error from small sample sizes. Included in the Spearman rank correlations was a fourth data set consisting of a compilation of data from experiments 2 and 3. This was done to increase the sample size obtained from these experiments, as they shared similar results and were conducted identically. The creation of the fourth combined dataset allowed for a better comparison of experiments 2 and 3 to experiment 1. In all Spearman rank correlation pairs there was a significant positive correlation found in the order of rank of the F_3 families (Table 3). In light of this, the data from experiment 1 was found to be acceptable for pooling with experiments 2 and 3 for the QTL mapping portion of this study (Chapter 3) (Figure 3). The continuous distribution of values in each of the experiments confirms the quantitative inheritance of parthenocarpic fruit set in cucumber, which will be explored in the following chapters.

The phenotypic evaluation used in this study can be compared to the methods used by de Ponti (1976) in cucumber, as well as Kikuchi et al. (2008) and Miyatake et al. (2012) in eggplant, since the number of pistillate flowers was strictly controlled at 25 per plant. De Ponti (1976) had proposed the use of a parthenocarpic percentage statistic as the most effective way to evaluate parthenocarpy in cucumber. The data recorded in this study can also be computed as parthenocarpic percentage on a per plant basis with the following formula (de Ponti, 1976):

% parthenocarpy = (number of parthenocarpic fruits/total number of pistillate flowers) x 100.

The use of this formula allows for more accurate evaluation of parthenocarpic expression when working with lines with differing numbers of pistillate flowers, differences in fruit size, and yield capacity. For simplification, parthenocarpic percentage was not formally used in this study as many of these factors were already accounted for in the selection of parent lines and the experimental treatment of plants.

Location of Parthenocarpic Fruit Initiation

The location of where fruit set occurs became important for a more detailed understanding of parthenocarpic fruit set. Each experiment was scored node by node along the main stem for the occurrence of parthenocarpic fruit initiation (Figure 4). The data revealed a difference in the average node of fruit initiation between experiment 1 and the other two experiments during the first 30 nodes of the plant (Figure 4). The difference appeared to reflect the observation that the lower nodes of some plants in experiment 1 were aborting flowers and caused what appears to be a delay in the onset of the fruit initiation.

In experiments 2 and 3, a bimodal distribution of fruit initiation is observed during the first 30 nodes of growth (Figure 4). This was anticipated to be due to source/sink relationships and reflects that once plants have begun fruit set they continue to set fruit until they are unable to support any additional fruits with available assimilates (Lee and Bazzaz, 1982a; Lee and Bazzaz, 1982b; Lloyd, 1980; Schapendonk and Brouwer, 1984; Stephenson, 1981; Stephenson et al., 1988; de Stigter, 1969). Following this, there appeared to be a quiescent period and many

flowers are aborted. Once active fruits reach a certain level of maturity, fruit set resumes and another flush of fruits is initiated. Experiment 1 is expected to follow this same phenomenon but because of the delayed fruit set, a bimodal distribution was not observed during the first 30 nodes of growth (Figure 4).

The bimodal distribution seen during the first 30 nodes of plant growth suggested that the ability of an individual plant to support a second flush of fruits may be a possible confounding factor to the accurate measurement of parthenocarpic fruit set. From the perspective of the application of parthenocarpy to processing cucumber varieties it should be noted that processing cucumber varieties rarely are allowed to reach 30 nodes in maturity during commercial production. These varieties are typically harvested for immature fruit when plants are approximately only 20 nodes in length (de Ponti, 1976) (Addendum 7). In consideration of this, a dataset comprised of data from node 6 to node 20 was created (Figure 5). Due to the complications in experiment 1 from delayed fruit set, this analysis was conducted using data only from experiments 2 and 3. The data from nodes 6 through 20 in experiments 2 and 3 were continuously distributed and resembled the distribution of data obtained when considering nodes 6 through 30, suggesting that effective phenotyping of parthenocarpic fruit set could be conducted on immature plants with only 20 nodes of plant growth (Figure 5).

Differences in the location of fruit set were observed between the parental lines in all three experiments (Figure 6). Throughout this study the highly parthenocarpic parent (‘2A’) consistently initiated fruit development very early on plant nodes 6 through 10 while the non-parthenocarpic parent (‘Gy8’) rarely initiated fruit development before node 10. It has been suggested that accurate preselection of young plants with superior parthenocarpic expression could be achieved by observing fruit set on the first five pistillate flowers of the plant by

breeding programs (de Ponti, 1976). Interested in the hypothesis that early parthenocarpic fruit set was indicative of overall parthenocarpic capacity, another dataset was created which was only inclusive of data collected from plant nodes 6 through 10 from experiments 2 and 3 (Figure 7). This dataset resembled a logarithmic distribution of values with the non-parthenocarpic parent ('Gy8') averaging nearly zero fruit per plant. Further discussion of early fruit set and the potential for preselection will be explored in Chapter 3.

Exploring the Effects of Plant Maintenance and Treatment

The plant maintenance and treatment experiment indicated that the pruning of lateral branches and the lower five nodes of each plant affects the number of parthenocarpic fruit initiated as well as the timing of fruit set (Table 4, Figure 8). Line '2A' showed no change in the number of parthenocarpic fruit initiated when comparing fruit initiation occurring on nodes 6 through 30 on both pruned and unpruned plants (Table 4). In contrast, 'Gy8' showed an increase in the number of parthenocarpic fruit initiated in pruned plants when considering plant nodes 6 through 30. Both lines yielded approximately 3.5 more initiated fruit per plant when flowers from lateral branches were included in the comparison (Table 4). However, these additional fruit on 'Gy8' were mostly set late and away from the crown of the plant. This differs from the pattern of fruit set in pruned 'Gy8' plants where fruit set occurs earlier (Figure 8). The increase in initiated parthenocarpic fruit on both parental lines when lateral branches are not disturbed is plausibly a direct result of increased plant photosynthetic capacity (Marcelis et al., 2004; Schapendonk and Brouwer, 1984; Stephenson, 1981). However, the early fruit set and increase in initiated fruit seen along the main stem of pruned 'Gy8' plants is potentially related to changes in hormone signaling and/or balance in response to wounding. Previous studies have shown that the removal

of apical and/or axillary shoot meristems promoted fruit growth and in some instances induced parthenocarpic fruit set (Carbonell and Garcia-Martinez, 1980; Coombe, 1962; Parry, 1976; Quinlan and Preston, 1971; Saunders et al., 1991; Serrani et al., 2010; Westwood and Bjornstad, 1974). In addition, dominance relationships between fruits and shoots have been demonstrated and it seems plausible that the removal of axillary shoots in this study disrupted the inhibition of fruit growth by the growing shoot and led to parthenocarpic fruit set (Bangerth, 1989; Gruber and Bangerth, 1990; Serrani et al., 2010; Westwood and Bjornstad, 1974). The timing of shoot removal has also been reported to result in increased fruit set and in some instances parthenocarpic fruit set has been observed when shoot removal occurs shortly before or after anthesis (Carbonell and Garcia-Martinez, 1980; Coombe, 1962; Quinlan and Preston, 1971; Saunders et al., 1991; Westwood and Bjornstad, 1974). Interestingly, the onset of parthenocarpic fruit set on pruned 'Gy8' plants after node 10 coincides with the timing of lateral shoot pruning in this study.

The implications of this for the larger greenhouse study are that the estimates of parthenocarpic potential for plants with little genetic potential for parthenocarpic fruit set were being over estimated. Plants with high genetic potential were presumably more accurately estimated as they set fruit until a maximum in plant load capacity was attained regardless of experimental treatment. Once this maximum plant load was attained, the plants failed to set any additional fruit until existing fruits matured and thus resulted in the observed bimodal fruit distributions. Although possibly confounded by the capacity of a plant to support multiple fruit, F_3 families with high genetic potential for parthenocarpic fruit set still scored as the highest yielding in the focus greenhouse study (Addendum 8). F_3 families with low genetic potential for parthenocarpic fruit set were the lowest yielding in the focus greenhouse study although these

families were observed with higher yields than would be expected for non-parthenocarpic lines due to responses to plant wounding (Addendum 8). This potentially explains why the non-parthenocarpic parent line 'Gy8' unexpectedly yielded multiple parthenocarpic fruit per plant in the focus study.

Future Focus

This study takes a new approach to assessing parthenocarpic fruit set potential. By focusing on fruit initiation and early fruit development, a major step has been taken in separating the true parthenocarpic character from yield related traits that have confounded past studies. This study may still be confounded by the capacity of individual plants to bear differing fruit loads. Future studies may wish to address this by instituting a continuous harvest of fruits as soon as fruits can be declared as either initiating development or failing to initiate development. However, following this approach may in itself be complicated by plant stresses and changes in fruit dominance if fruits are being continuously removed from the plant (Gruber and Bangerth, 1990). This exemplifies the complexity in accurately assessing parthenocarpic potential. Though an idealized protocol may not be obtainable for accurately phenotyping parthenocarpic potential, future studies should continue to focus on early fruit development as the key to parthenocarpic fruit set.

Literature Cited

Aalbersberg, W. and G. van Wijchen. 1987. Parthenocarpic gherkins: newcomer is a distinct improvement. Groenten En Fruit. 42:66-67.

Baker, L.R., J.W. Scott, and J.E. Wilson. 1973. Seedless pickles - a new concept. Res. Rep.

Mich. St. Univ. agric. Exp. Stn. 227:1-12.

Bangerth, F. 1989. Dominance among fruit/sinks and the search for a correlative signal. Physiol. Plant. 76:608-614.

Bartlett, M. S. 1937. Properties of sufficiency and statistical tests. Proc. R. Soc. A. 160:268-282.

Beyer, E.M. Jr. and B. Quebedeaux. 1974. Parthenocarpy in cucumber: mechanism of action of auxin transport inhibitors. J. Amer. Soc. Hort. Sci. 99:385-390.

Cantliffe, D.J. and S.C. Phatak. 1975. Use of ethephon and chlorflurenol in once-over pickling cucumber production system. J. Amer. Soc. Hort. Sci. 100:264-267.

Carbonell, J. and J.L. Garcia-Martinez. 1980. Fruit-set of unpollinated ovaries of *Pisium sativum* L.: influence of vegetative parts. Planta. 147:444-450.

Choudhury, B. and S.C. Phatak. 1959. Sex expression and fruit development in cucumber (*Cucumis sativus* L.) as affected by gibberellin. Indian J. Hortic. 16:233-235.

Coombe, B.G. 1962. The effect of removing leaves, flowers and shoot tips on fruit-set in *Vitis vinifera* L. J. Hortic. Sci. 37:1-15.

Denna, D.W. 1973. Effect of genetic parthenocarpy and gynoecious habit on fruit production and growth of cucumber, *Cucumis sativus* L. J. Amer. Soc. Hort. Sci. 98:602-604.

Elassar, G., J. Rudich, D. Palevitch, and N. Kedar. 1974. Induction of parthenocarpic fruit development in cucumber by growth regulators. HortScience. 9:238-239.

El-Shawaf, I.I.S. and L.R. Baker. 1981. Inheritance of parthenocarpic yield in gynoecious pickling cucumber for once-over mechanical harvest by diallel analysis of six gynoecious lines. J. Amer. Soc. Hort. Sci. 106:365-370.

"FAOSTAT." FAOSTAT. N.p., 08 Aug. 2013. Web. 19 Sept. 2013. <http://faostat.fao.org/>.

Fos, M., F. Nuez, and J.L. Garcıa-Martınez. 2000. The gene pat-2, which induces natural parthenocarpy, alters the gibberellin content in unpollinated tomato ovaries. Plant Physiol. 122:471-480.

Fu, F.Q., W.H. Mao, K. Shi, Y.H. Zhou, T. Asami, and J.Q. Yu. 2008. A role of brassinosteroids in early fruit development in cucumber. J. Exp. Bot. 59:2299-2308.

Gillaspy, G., H. Ben-David, and W. Gruissem. 1993. Fruits: a developmental perspective. Plant Cell. 5:1439-1451.

Gruber, J. and F. Bangerth. 1990. Diffusible IAA and dominance phenomena in fruits of apple and tomato. Physiol. Plant. 79:354-358.

Gusmini, G. and T.C. Wehner. 2008 Fifty-five years of yield improvement for cucumber, melon, and watermelon in the United States. HortTechnology. 18:9-12.

Hawthorn, L.R. and R. Wellington. 1930. Geneva, a greenhouse cucumber that develops fruit without pollination. NY (Geneva) Agric Exp. Stat. Bull. 580:223-234.

Homan, D. 1964. Auxin transport in physiology of fruit development. Plant Physiol. 39:982–986.

Kikuchi, K., I. Honda, S. Matsuo, M. Fukuda, and T. Saito. 2008. Stability of fruit set of newly selected parthenocarpic eggplant lines. Sci. Hort. 115:111-116.

Kim, I.S., H. Okubo, and K. Fujieda. 1992. Endogenous levels of IAA in relation to parthenocarpy in cucumber (*Cucumis sativus* L.). Sci Hort. 52:1-8.

Kvasnikov, B.V., N.T. Rogova, S.I. Tarakonova, and I. Ignatova. 1970. Methods of breeding vegetable crops under the covered ground. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 42:45-57.

Lee, T.D. and F.A. Bazzaz. 1982a. Regulation of fruit and seed production in an annual legume, *Cassia fasciculata*. Ecology. 63:1363-1373.

Lee, T.D. and F.A. Bazzaz. 1982b. Regulation of fruit maturation pattern in an annual legume, *Cassia fasciculata*. Ecology. 63:1374-1388.

Lloyd, D.G. 1980. Sexual strategies in plants. I. An hypothesis on serial adjustment of maternal investment during one reproductive session. New Phytol. 86:69-79.

Marcelis, L.F.M., E. Heuvelink, L.R. Baan Hofman-Eijer, J. Den Bakker, and L.B. Xue. 2004. Flower and fruit abortion in sweet pepper in relation to source and sink strength. J. Exp. Bot. 55:2261-2268.

Meshcherov, E.T. and L.W. Juldasheva. 1974. Parthenocarpy in cucumber. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 51:204-213.

McCollum, J.P. 1934. Vegetable and reproductive responses associated with fruit development in cucumber. Mem. Cornell Univ. agric. Exp. Stn 163:27pp.

Miyatake, K., T. Saito, S. Negoro, H. Yamaguchi, T. Nunome, A. Ohyama, and H. Fukuoka. 2012. Development of selective markers linked to a major QTL for parthenocarpy in eggplant (*Solanum melongena* L.). Theor. Appl. Genet. 124:1403-1413.

Molesini, B., T. Pandolfini, G.L. Rotino, V. Dani, and A. Spena. 2009. Aucsia gene silencing causes parthenocarpic fruit development in tomato. Plant Physiol. 149:534-548.

Nijs, A.P.M.D. and D.L. Visser. 1980. Induction of male flowering in gynoecious cucumbers

(*Cucumis sativus* L.) by silver ions. Euphytica. 29:273-280.

Pandolfini, T. 2009. Seedless fruit production by hormonal regulation of fruit set. Nutrients. 1:168-177.

Parry, M.S. 1976. Effect of pruning and partial disbudding on the cropping of Cornice and Beurre Hardy pears. J. Hortic. Sci. 51:159-166.

Pascual. L., J.M. Blanca, J. Canizares, and F. Nuez. 2009. Transcriptomic analysis of tomato carpel development reveals alterations in ethylene and gibberellins synthesis during *pat3/pat4* parthenocarpic fruit set. BMC Plant Biol. 9:67.

Pike, L.M. and C.E. Peterson. 1969. Inheritance of parthenocarpy in the cucumber (*Cucumis sativus* L.) Euphytica. 18:101-105.

de Ponti, O.M.B. 1976. Breeding parthenocarpic pickling cucumbers (*Cucumis sativus* L): necessity, genetical possibilities, environmental influences and selection criteria. Euphytica. 25:29-40.

de Ponti, O.M.B. and F. Garretsen. 1976. Inheritance of parthenocarpy in pickling cucumbers (*Cucumis sativus* L.) and linkage with other characters. Euphytica. 25:633-642.

Quinlan, J.D. and A.P. Preston. 1971. The influence of shoot competition on fruit retention and cropping of apple trees. J. Hortic. Sci. 46:525-534.

Robinson, R.W., D.J. Cantliffe, and S. Shannon. 1971. Morphactin-induced parthenocarpy in cucumber. Science. 171:1251-1252.

Robinson, R.W. and S. Reiners. 1999. Parthenocarpy in summer squash. HortScience. 34(4):715-717.

Rotino, G.L., E. Perri, M. Zottini, H. Sommer, and A. Spena. 1997. Genetic engineering of parthenocarpic plants. Nat. Biotechnol. 15:1398-1401.

Ruan, Y.L., J.W. Patrick, M. Bouzayen, S. Osorio, and A.R. Fernie. 2012. Molecular regulation of seed and fruit set. Trends Plant Sci. 17:656-665.

Saunders, R.C., G. Jacobs, and D.K. Strydom. 1991. Effect of pruning on fruitset and shoot growth of ‘Packham’s Triumph’ pear trees. Sci. Hort. 47:239-245.

Schapendonk, A.H.C.M. and P. Brouwer. 1984. Fruit growth of cucumber in relation to assimilate supply and sink activity. Sci. Hort. 23:21-33.

Serrani, J.C., E. Carrera, O. Ruiz-Rivero, L. Gallego-Giraldo, L.E.P. Peres, and J.L. Garcia-Martinez. 2010. Inhibition of auxin transport from the ovary or from the apical shoot induces parthenocarpic fruit-set in tomato mediated by gibberellins. Plant Physiol.

153:851-862.

Spearman, C. 1904. The proof and measurement of association between two things. Am. J. Psychol. 15:72-101.

Stephenson, A.G. 1981. Flower and fruit abortion: proximate causes and ultimate functions. Annu. Rev. Ecol. Syst. 12:253-279.

Stephenson, A.G., B. Delvin, and J.B. Horton. 1988. The effects of seed number and prior fruit dominance on the pattern of fruit production in *Cucurbita pepo* (Zucchini Squash). Ann. Botany. 62:653-661.

de Stigter, H.C.M. 1969. Growth relations between individual fruits, and between fruits and roots in cucumber. J. Exp. Bot. 27:87-97.

Strong, W.J. 1921. Greenhouse cucumber breeding. Proc. Am. Soc. Hort. Sci. 18:271-273.

Sturtevant, E.L. 1890. Seedless fruits. Mem. Torrey Bot. Club. 1:141-185.

Sun, Z. 2004. Inheritance and molecular marker-based genetic mapping of parthenocarpy in cucumber (*Cucumis sativus* L.). PhD. Thesis, University of Wisconsin-Madison.

Sun, Z., R.L. Lower, and J.E. Staub. 2006a. Variance component analysis of parthenocarpy in elite U.S. processing type cucumber (*Cucumis sativus* L.) lines. Euphytica. 148:331-339.

Sun, Z., J.E. Staub, S.M. Chung, and R.L. Lower. 2006b. Identification and comparative analysis of quantitative trait loci (QTL) associated with parthenocarpy in processing cucumber. Plant Breeding. 125:281-287.

Tiedjens, V.A. 1928. Sex ratios in cucumber flowers as affected by different conditions of soil and light. J. agric. Res. 36:721-746.

Uzcategui, N.A. and L.R. Baker. 1979. Effects on multiple pistillate flowering on yields of gynoecious pickle cucumbers. J. Amer. Soc. Hort. Sci. 104:148-151.

Varoquaux, F., R. Blanvillain, M. Delseny, and P. Gallois. 2000. Less is better: new approaches for seedless fruit production. Trends Biotechnol. 18:233-239.

Westwood, M.N. and H.O. Bjornstad. 1974. Fruitset as related to girdling, early cluster thinning and pruning of 'Anjou' and 'Comice' pear. HortScience. 9:342-344.

Yin, Z., R. Malinowski, A. Ziolkowska, H. Sommer, W. Plader, and S. Malepszy. 2006. The DefH9-iaaM-containing construct efficiently induces parthenocarpy in cucumber. Cell. Mol. Biol. Lett. 11:279-290.

Table 1. The number of fruit initiated per plant in each of the three greenhouse experiments conducted to study parthenocarpic fruit set in a 2A×GY8 $F_{2:3}$ population of *C. sativus*.

Exp	Mean	95% Mean CI[z]	Std Dev	Median	Max	Min	Total Plants[y]
Exp 1	3.23	3.10-3.36	2.56	3	15	0	1018
Exp 2	3.74	3.57-3.90	1.71	4	10	0	653
Exp 3	3.76	3.60-3.92	1.70	4	11	0	658

[z]95% Mean Confidence Interval.

[y]Experiment 1 contained 5 F_3 plants from each of 201 F_3 families. Experiments 2 and 3 contained 3 F_3 plants from each of 205 F_3 families. All Experiments contained 20 plants of each parental line and 5 2A×GY8 F_1.

Table 2. The p-values obtained from two sample t-tests used for comparing of each of the three greenhouse experiments conducted to study parthenocarpic fruit set in a 2A×Gy8 $F_{2:3}$ population of *C. sativus*.

	Exp 1	Exp 2	Exp 3
Exp 1		8.43E-06***	2.91E-06***
Exp 2	8.43E-06***		0.80
Exp 3	2.91E-06***	0.80	

***Calculated values were found to be significant at alpha = 0.01.

Table 3. Spearman's rank correlation coefficients (Spearman, 1904) from comparisons of F_3 family means in each of the three greenhouse experiments conducted to study parthenocarpic fruit set in a 2A×Gy8 $F_{2:3}$ population of *C. sativus*. Due to changes in experimental design between experiment 1 and experiments 2 and 3, which were conducted identically, rho values are also presented comparing experiment 1 with a data set consisting of combined data from experiments 2 and 3.

	Exp 1	Exp 2	Exp 3	Exps 2 and 3[z]
Exp 1		0.48***	0.49***	0.56***
Exp 2	0.48***		0.54***	X[y]
Exp 3	0.49***	0.54***		X[y]
Exps 2 and 3[z]	0.56***	X[y]	X[y]	

[z]Combined data set containing the combined data of experiments 2 and 3.

[y]Rho values were not calculated for comparisons of experiment 2 or experiment 3 to the combined data set containing data from both experiments 2 and 3.

***Calculated values were found to be significant at alpha = 0.01.

Table 4. Comparisons of the number of fruit initiated on plants of inbred lines '2A' and 'Gy8' under different plant maintenance regimens. Pruned plants were cleared of all lateral branches, foliage, and flowers from the bottom five nodes of the plant. Unpruned plants were not pruned during growth and allowed to grow unhindered. The Unpruned With Laterals and Unpruned datasets utilize the same plants but are evaluated under two different protocols. The Pruned dataset is collected from separate plants.

UNPRUNED WITH LATERALS		UNPRUNED		PRUNED	
ENTRY	FRUIT INITIATED	ENTRY	FRUIT INITIATED	ENTRY	FRUIT INITIATED
2A-1	10	2A-1	7	2A-1	6
2A-2	11	2A-2	5	2A-2	9
2A-3	8	2A-3	6	2A-3	5
2A-4	9	2A-4	5	2A-4	5
2A-5	8	2A-5	6	2A-5	4
AVERAGE	9.2	AVERAGE	5.8	AVERAGE	5.8
ENTRY	FRUIT INITIATED	ENTRY	FRUIT INITIATED	ENTRY	FRUIT INITIATED
Gy8-1	5	Gy8-1	0	Gy8-1	2
Gy8-2	3	Gy8-2	1	Gy8-2	2
Gy8-3	2	Gy8-3	1	Gy8-3	3
Gy8-4	5	Gy8-4	0	Gy8-4	3
Gy8-5	5	Gy8-5	0	Gy8-5	3
AVERAGE	4	AVERAGE	0.4	AVERAGE	2.6

Figure 1. A: Photograph illustrating the differences between dried unpollinated *C. sativus* ovaries that have initiated parthenocarpic fruit development prior to aborting (left) and dried unpollinated ovaries that aborted at anthesis (right). At center are unpollinated ovaries at anthesis for comparison. B: Photograph illustrating examples of parthenocarpic fruit that have aborted after multiple days of fruit development.

A.

B.

Figure 2. Frequency distributions of the total number of fruit initiated per plant in experiments 1-3. Data was collected from plant nodes 6 thru 30 for a maximum possible total of 25 initiated fruit per plant. In each experiment the average values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Experiment 1 consisted of 1050 2A×Gy8 F_3 and accompanying control plants. B: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants. C: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants.

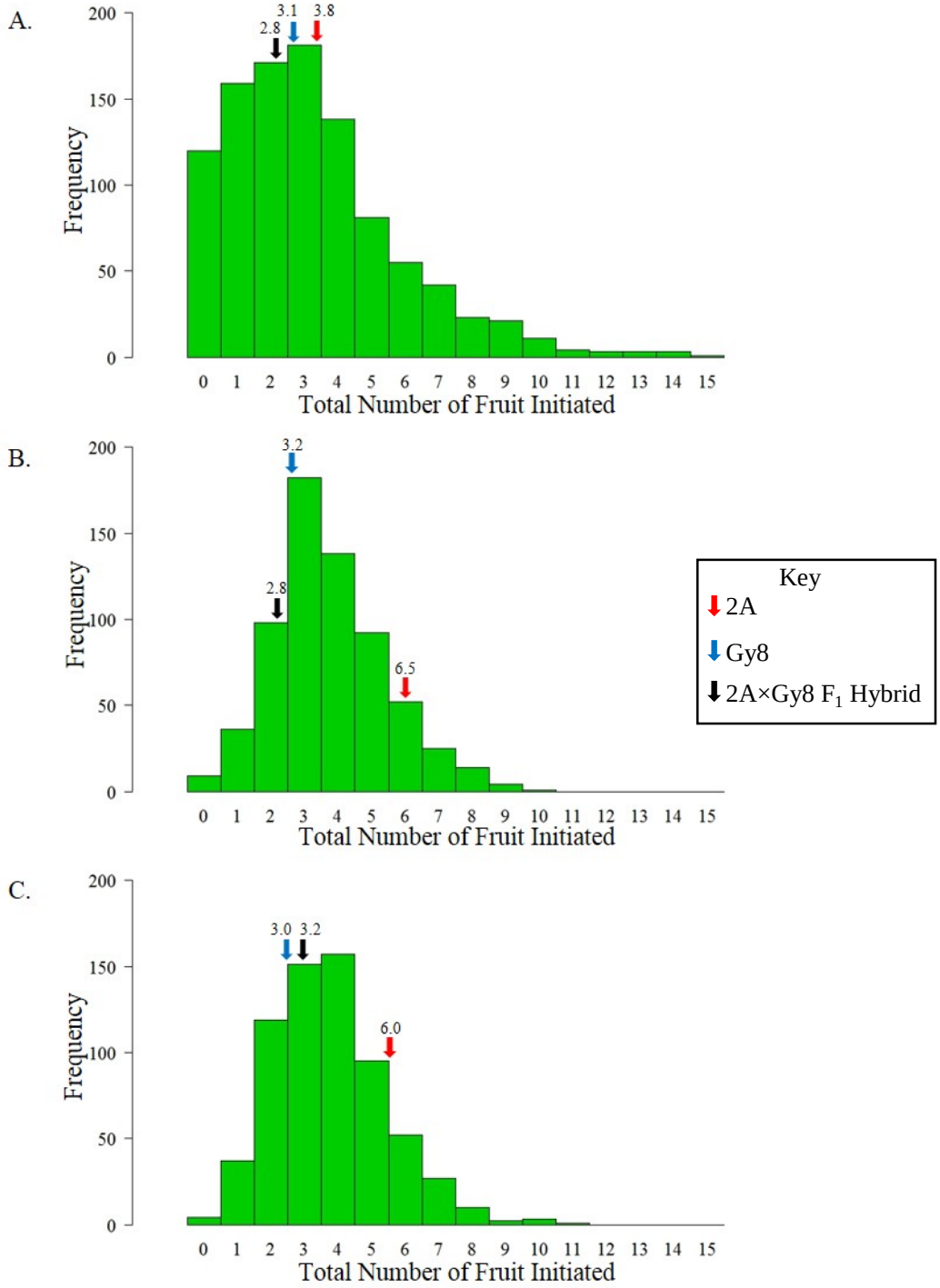


Figure 3. Frequency distributions of the total number of fruit initiated per plant in the pooling of experiments 1-3. Data was collected from plant nodes 6 thru 30 for a maximum possible total of 25 initiated fruit per plant. In each figure the average pooled values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Frequency distribution of the total number of fruit initiated per plant in the pooling of experiments 1-3. The pooled experiments together consisted of 2370 2A×Gy8 F_3 and accompanying control plants. B: Frequency distribution of the average number of fruit initiated for 205 F_3 families obtained by pooling across the three experiments. Each F_3 family is represented by 11 F_3 plants.

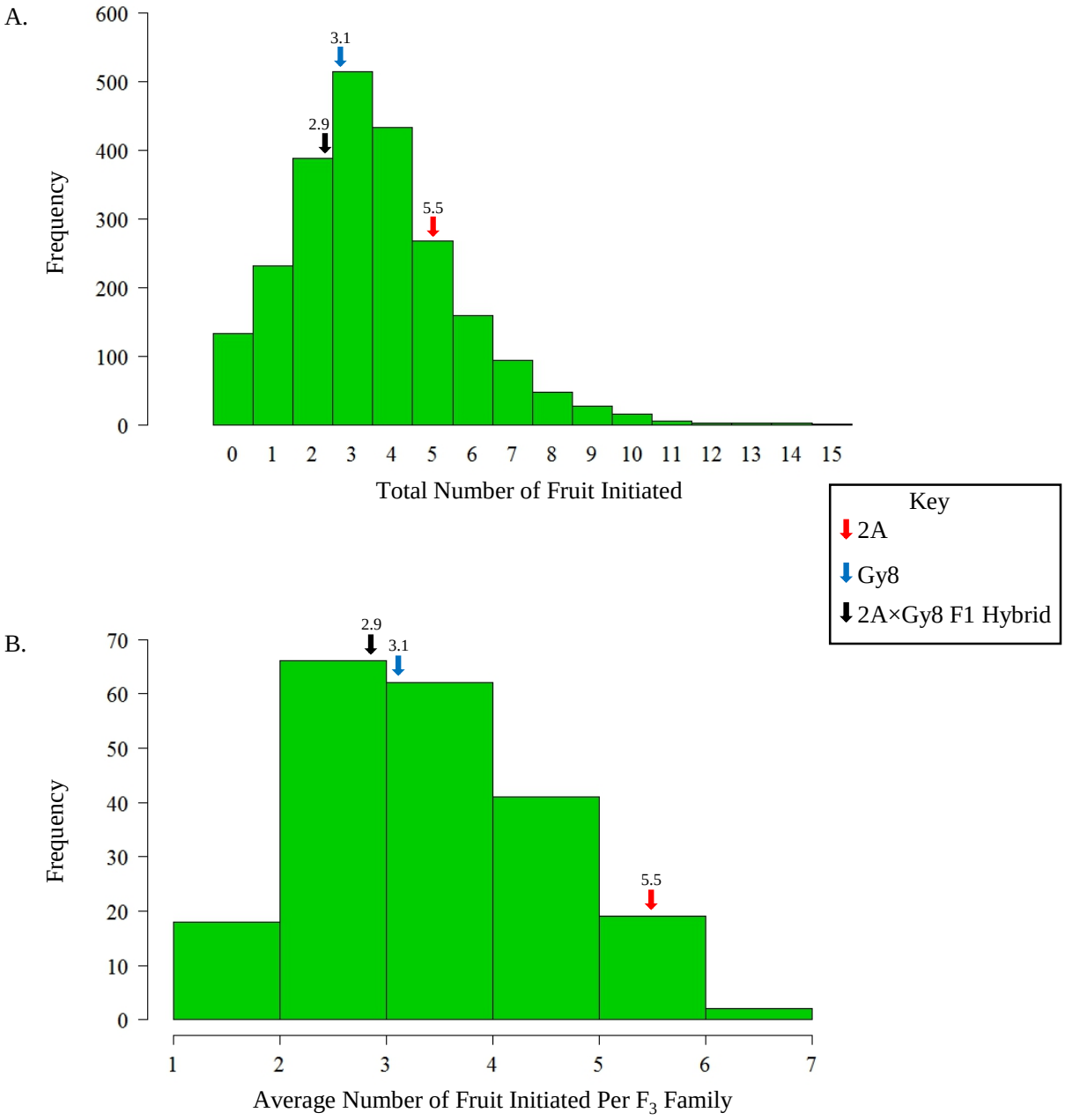


Figure 4. Frequency distributions displaying the frequency of fruit set initiation at each plant node across experiments 1-3. All plants were cleared of flowers and vegetation on nodes 1 thru 5. A: Experiment 1 consisted of 1050 2A×Gy8 F_3 and accompanying control plants. B: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants. C: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants.

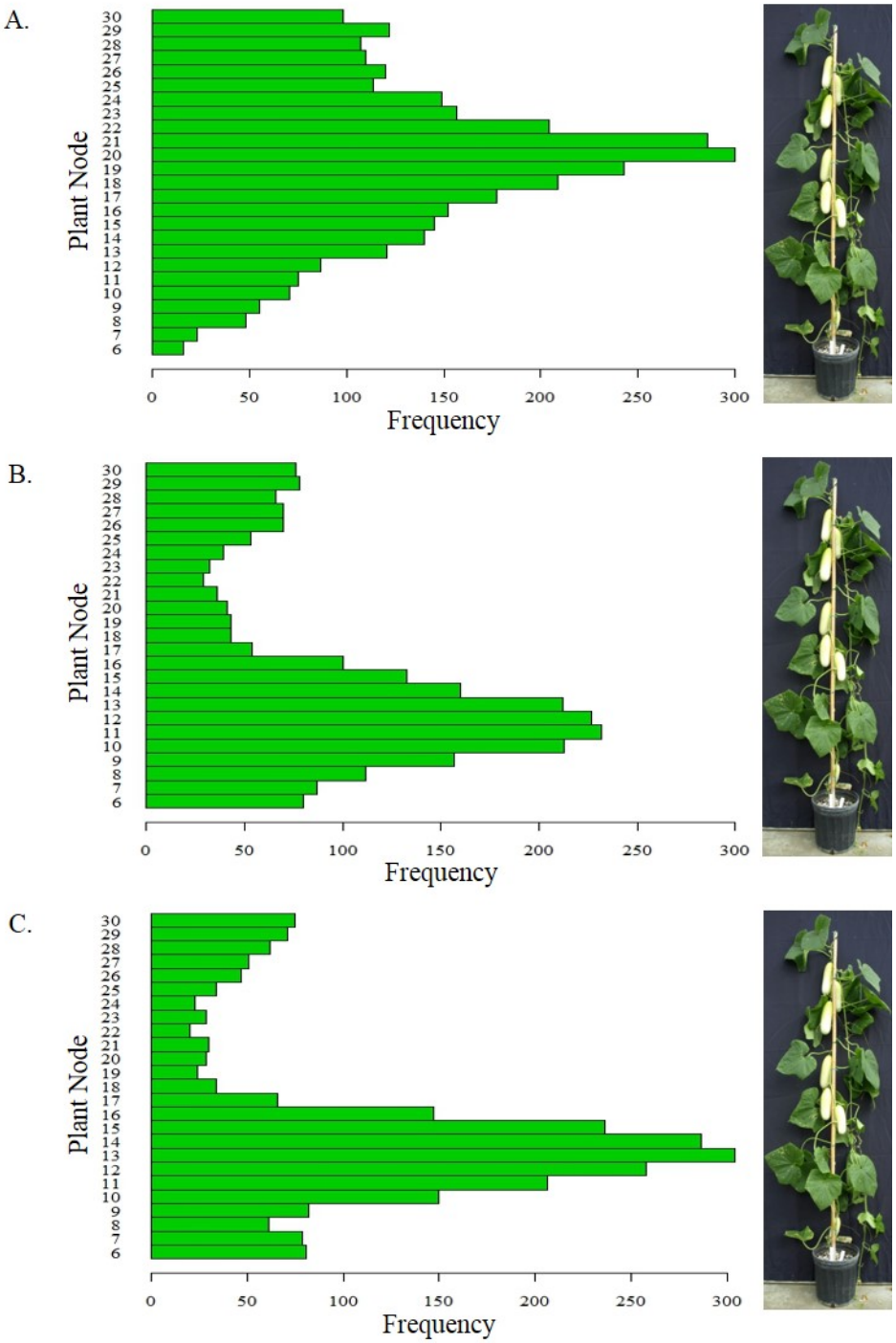


Figure 5. Frequency distributions of the total number of fruit initiated per plant in experiments 2 and 3. Data was collected from plant nodes 6 thru 20 for a maximum possible total of 15 initiated fruit per plant. In each experiment the average values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants and was grown from September 2011 to December 2011. B: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants and was grown from March 2012 to June 2012.

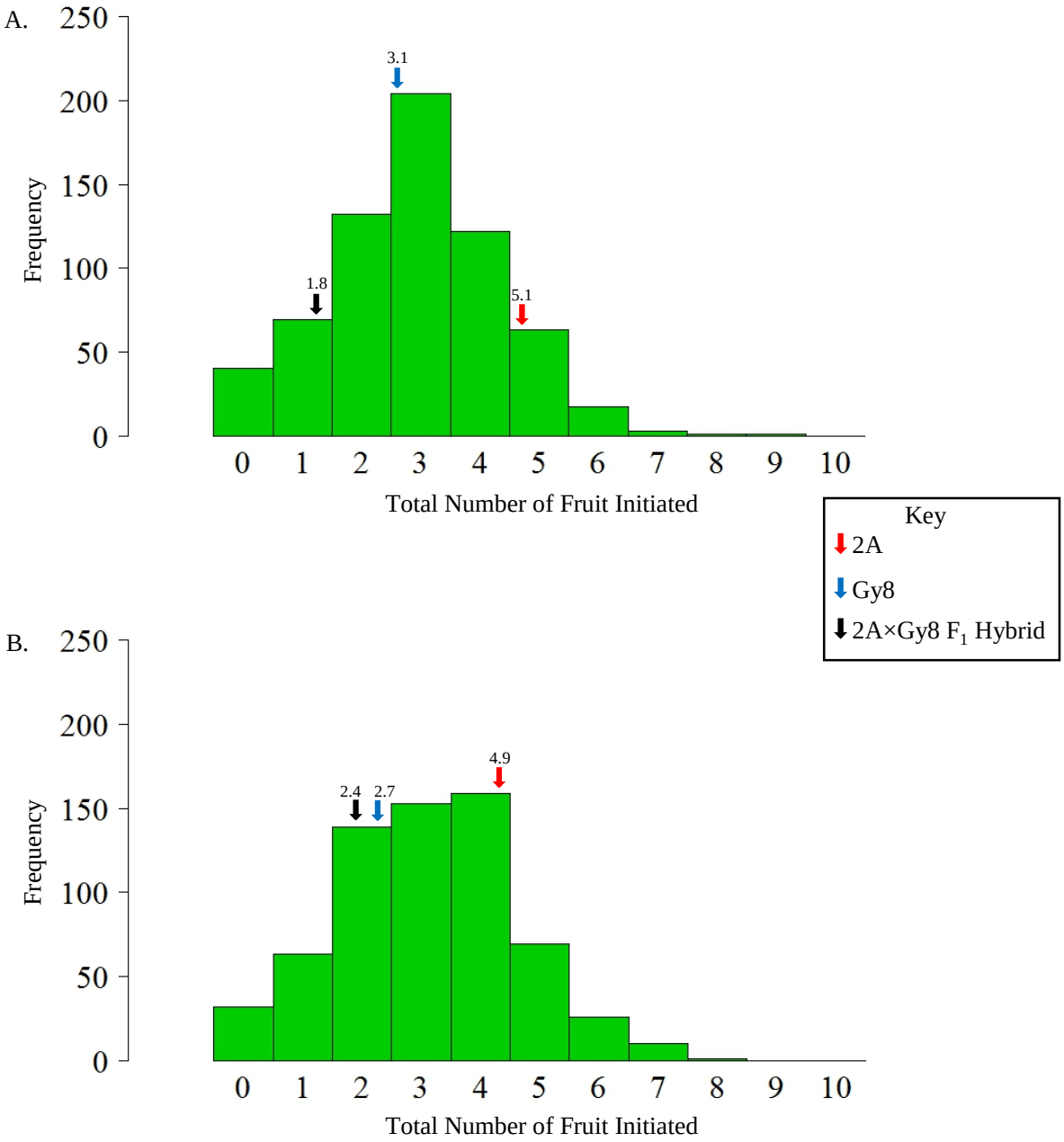


Figure 6. Photograph depicting the typical fruit number and fruit set location in *C. sativus* parental lines ‘2A’ (left) and ‘Gy8’ (right) and the 2A×Gy8 F_1 hybrid (center).

Figure 7. Frequency distributions of the total number of fruit initiated per plant in experiments 2 and 3. Data was collected from plant nodes 6 thru 10 for a maximum possible total of five initiated fruit per plant. In each experiment the average values of the control parental lines and 2A×Gy8 hybrid are designated with arrows. A: Experiment 2 consisted of 660 2A×Gy8 F_3 and accompanying control plants. B: Experiment 3 consisted of 660 2A×Gy8 F_3 and accompanying control plants.

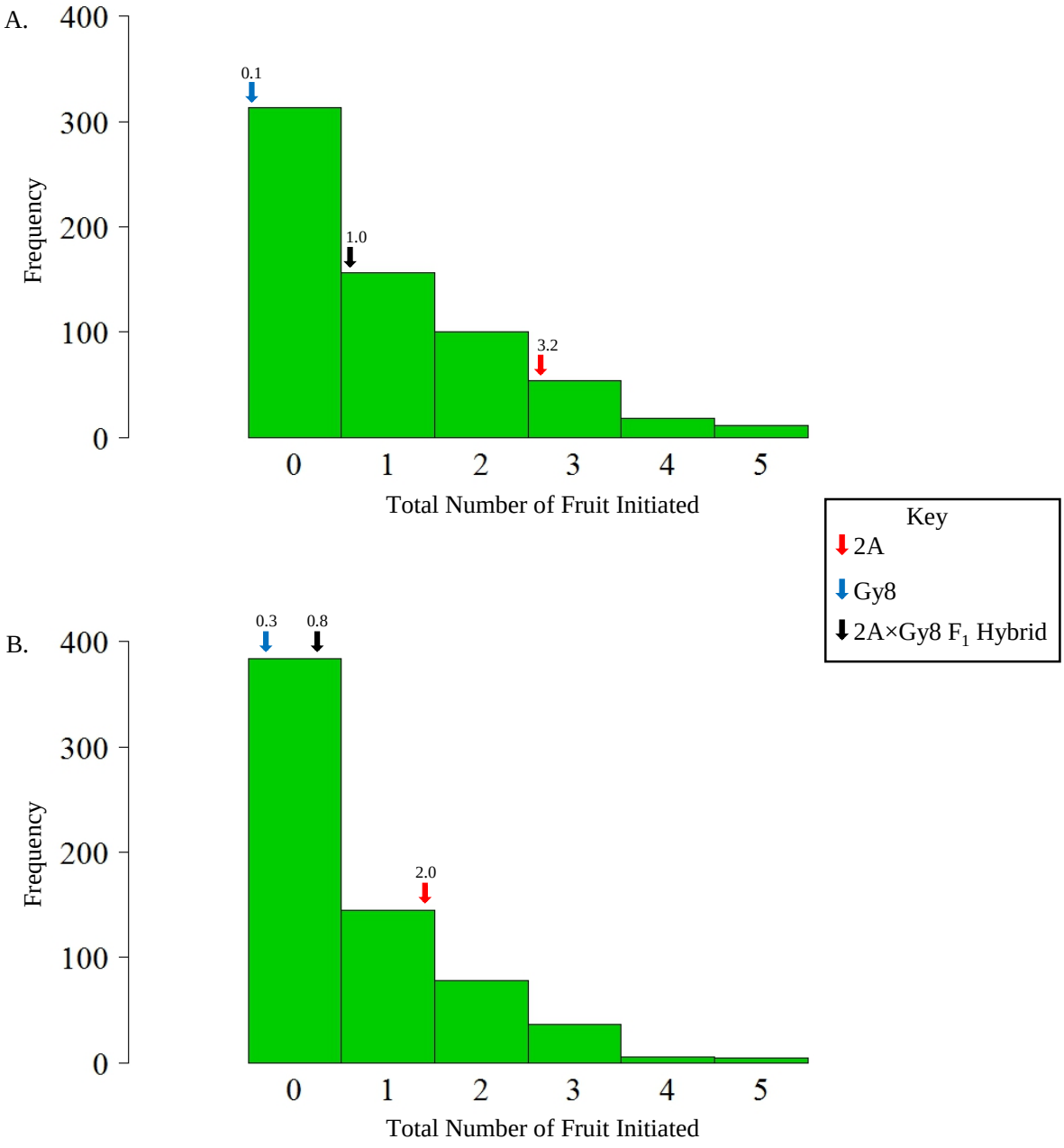


Figure 8. Photographs of ‘2A’ (A) and ‘Gy8’ (B) under pruned and unpruned treatments in environmentally controlled greenhouse conditions. In each photograph the plant on the left was not trimmed of lateral branches or cleared of foliage from the bottom five nodes of the plant. The plant on the right was trimmed of all lateral branches routinely throughout growth and was cleared of all foliage from the bottom five nodes of the plant.

A.

B.

Addendum 1. Descriptions of the origin and trait characteristics of inbred lines ‘2A’, ‘Gy8’, and ‘Gy7’.

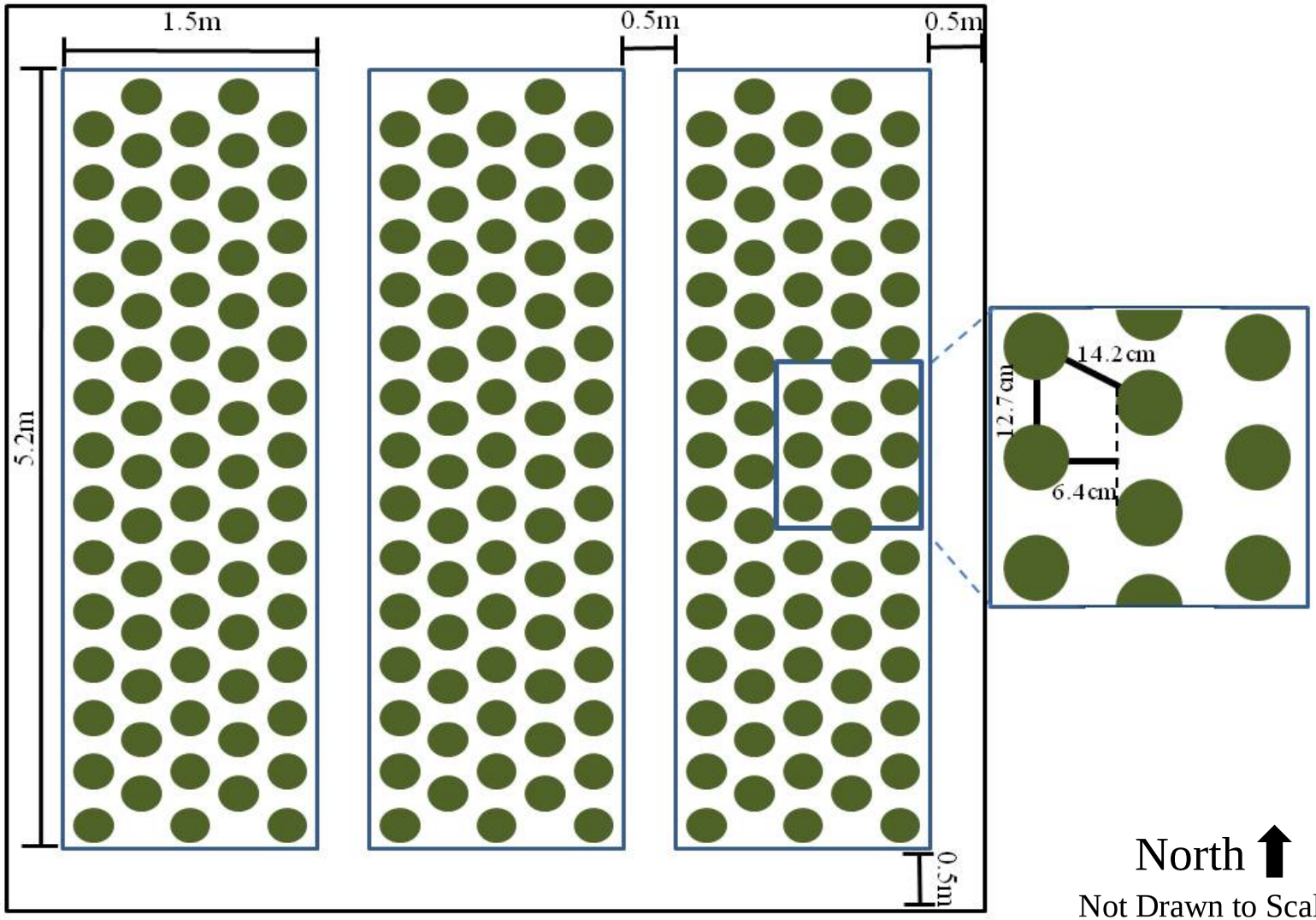
Inbred line ‘2A’ originated from an attempt to identify genetic sources with stronger expression of gynoecy and to combine it with disease resistance, high yield, and improved fruit quality in cucumber. A population designated Gynoecious Synthetic (GS; not publicly released) was developed from the random mating of 50 gynoecious lines and hybrids. The GS population was subjected to more than 10 generations of half-sib selection for gynoecious sex expression and high fruit number per plant in both pollinated and pollen free field plantings. A highly parthenocarpic line developed from the GS population was selected and crossed with the gynoecious line ‘Gy7’ in order to improve disease resistance, fruit quality, and horticultural characteristics of the new line. This line was designated ‘2A’. The ‘2A’ inbred used in this study is a F_9 selection from this cross. ‘2A’ is indeterminate, gynoecious, and has stable parthenocarpic expression in a wide range of environments. Fruits have warts with stippling. ‘2A’ is resistant to scab, cucumber mosaic virus, downy mildew, anthracnose, angular leaf spot, and has some field tolerance to powdery mildew. (Sun, 2004).

Inbred line ‘Gy8’ is an indeterminate and gynoecious advanced selection derived from a cross of processing cucumber lines ‘Gy14A’ and ‘UW70’. ‘UW70’ has a complex pedigree including contributions from lines: ‘MSU 713-5’, ‘MSU 7’, ‘New Hampshire PM #1 Bush’, ‘New Hampshire Tiny Dill’, ‘Chipper’, and ‘SC 10’. ‘Gy8’ has moderately long, medium green vines and an indeterminate, branched habit similar to ‘Gy14A’. Fruits are cylindrical with slightly rounded to blocky ends, light to medium green color, white spines, moderate warts, moderate stippling, and moderate striping. Fruits of ‘Gy8’ are generally about 0.3 of an L/D unit longer than those of ‘Gy14A’. ‘Gy8’ is resistant to scab, cucumber mosaic virus, downy mildew, anthracnose, angular leaf spot, and has some field tolerance to powdery mildew under

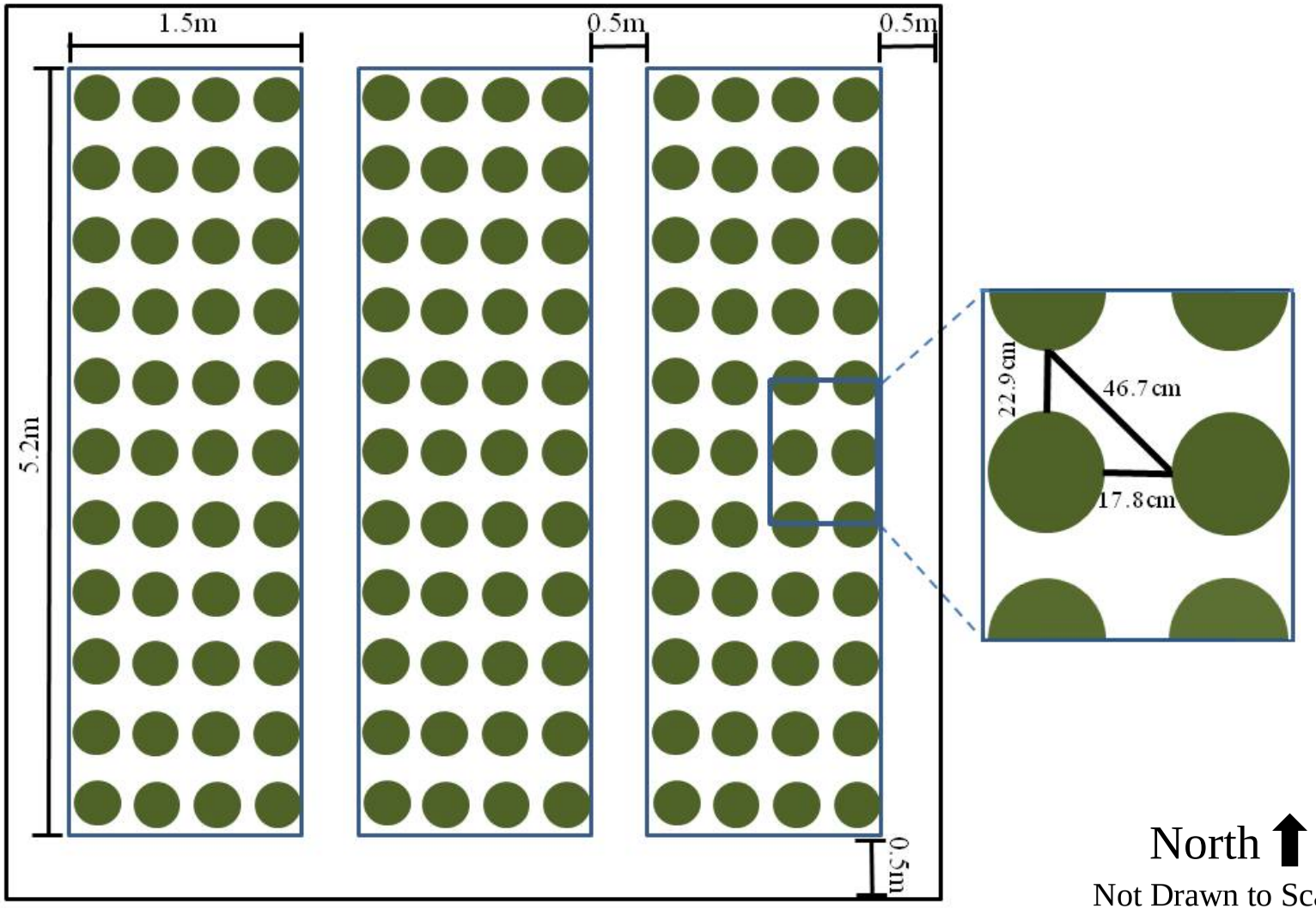
Wisconsin conditions. ‘Gy8’ has good combining ability for fruit number and hybrids with ‘Gy8’ parentage have performed well in trials located in all major processing cucumber production areas in the United States (Lower, 1996).

Inbred line ‘Gy7’ is a determinate, gynoecious advanced selection derived from a cross of processing cucumber lines ‘Gy4’ and ‘M21’. ‘Gy7’ has dark green vines with one to five laterals. Fruits of ‘Gy7’ are longer than ‘Gy14A’ with tapered ends, white spines, dark green medium size warts, slight stippling, and slight striping. ‘Gy7’ is resistant to scab, downy mildew, anthracnose and angular leaf spot under Wisconsin conditions. ‘Gy7’ has good combining ability for fruit number and hybrids with ‘Gy7’ parentage have performed well in trials located in all major processing cucumber production areas in the United States (Lower, 1996).

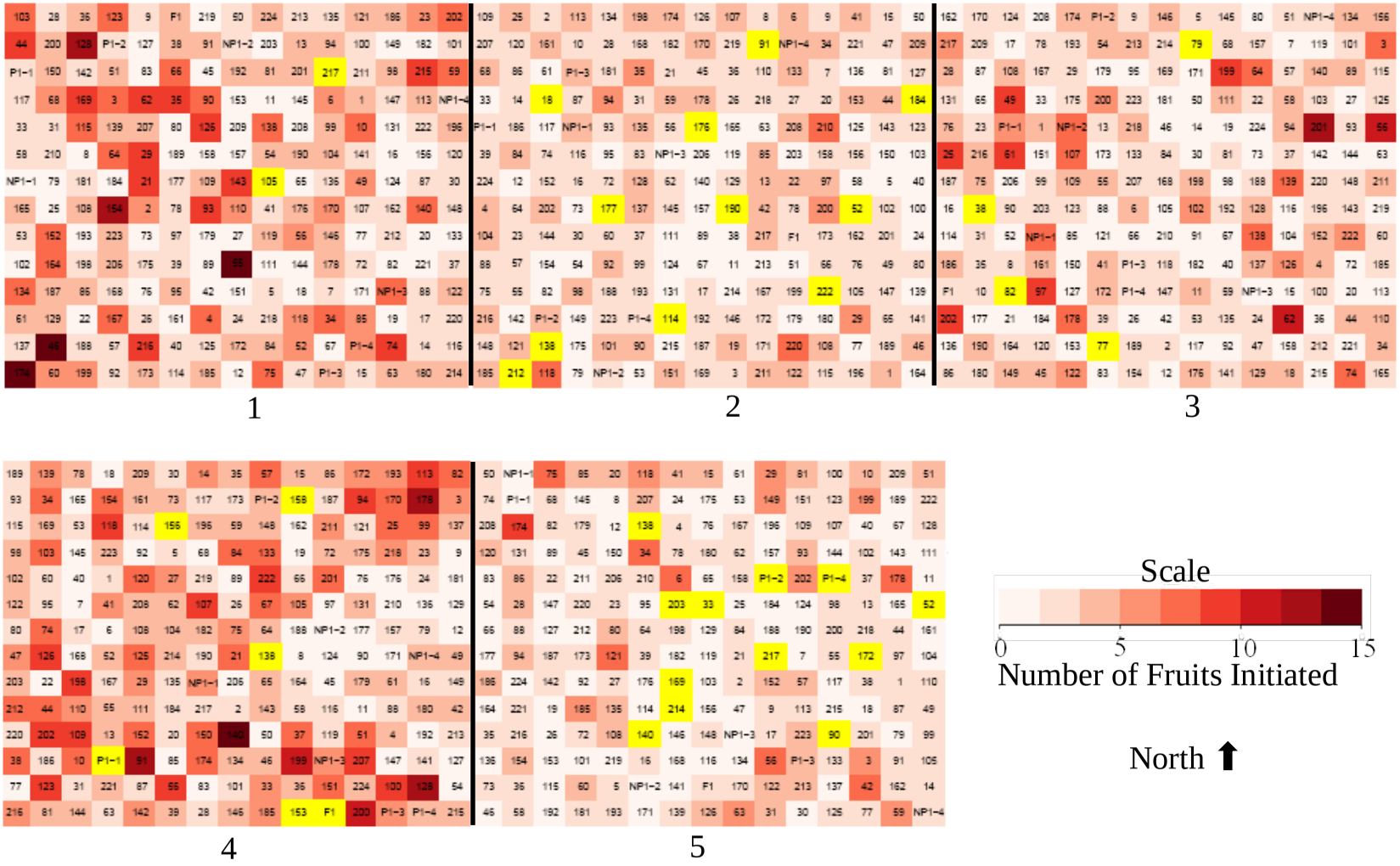
Addendum 2. Schematic for the plant spacing and layout in each of the five greenhouses utilized in experiment 1 at the University of Wisconsin-Madison Walnut Street Research Greenhouses located in Madison, WI. Experiment 1 consisted of 1050 2A×Gy8 F_3 and accompanying control plants. Each greenhouse in experiment 1 contained 1 F_3 plant from each of 201 2A×Gy8 F_3 families and 9 control plants for a total of 210 plants.



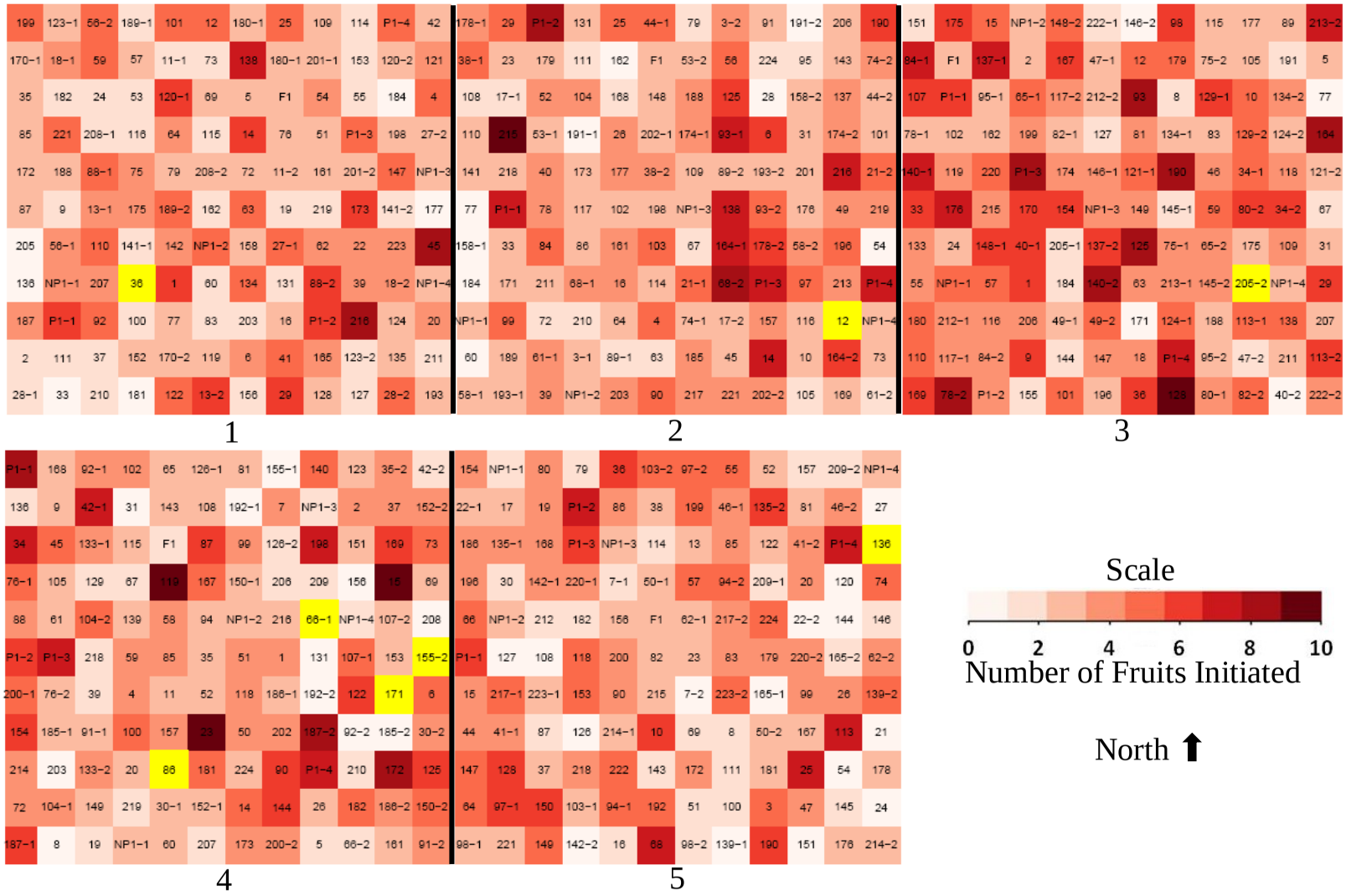
Addendum 3. Schematic for the plant spacing and layout in each of the five greenhouses utilized in experiments 2 and 3 at the University of Wisconsin-Madison Walnut Street Research Greenhouses located in Madison, WI. Experiments 2 and 3 each consisted of 660 2A×Gy8 F_3 and accompanying control plants. In experiments 2 and 3, three plants from each of 205 2A×Gy8 F_3 families were randomly distributed across each experiment. Each greenhouse contained 123 randomized 2A×Gy8 F_3 plants and 9 control plants for a total of 132 plants.



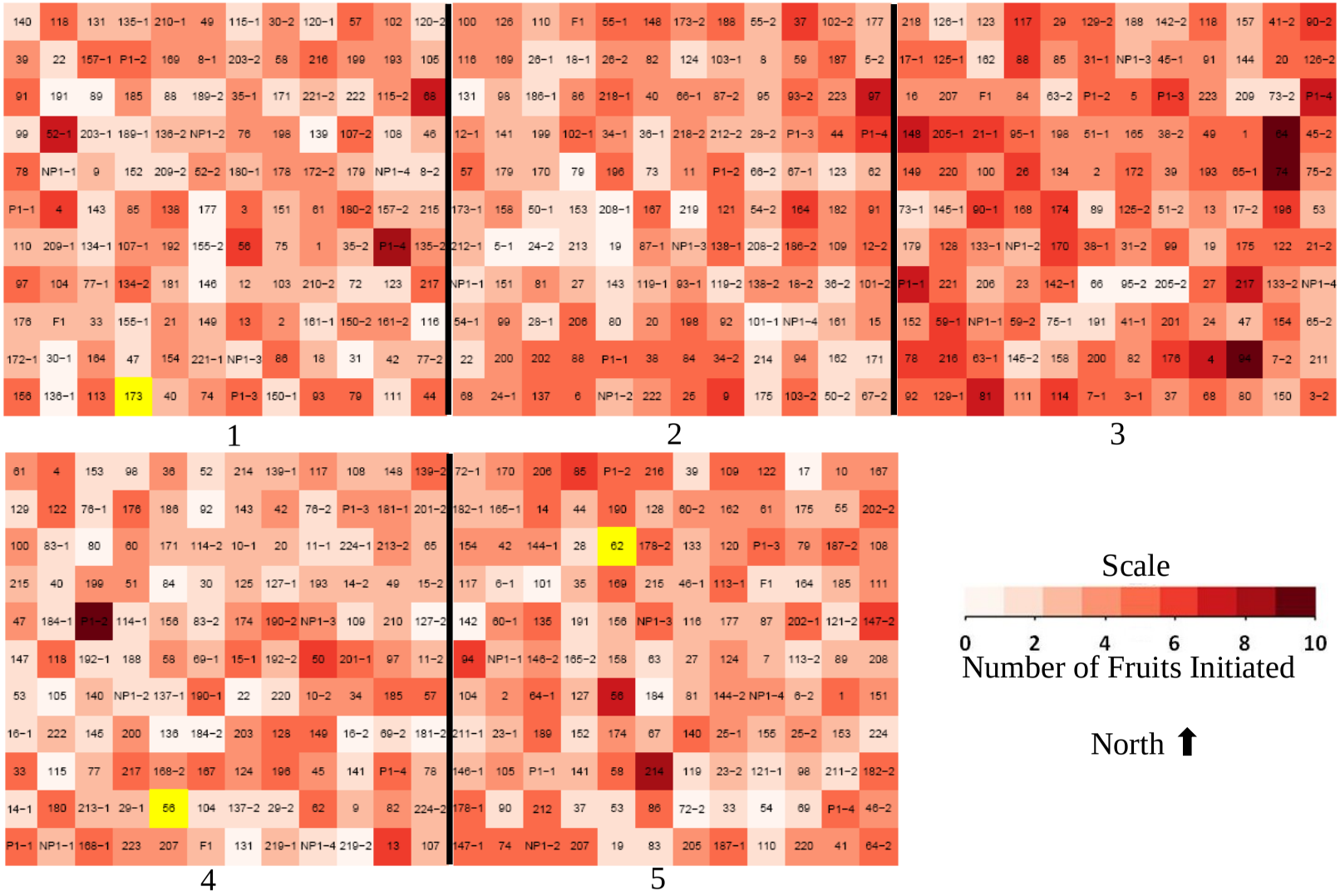
Addendum 4. Heat map displaying the spatial position and the number of initiated fruits set on each experimental plant during experiment 1 in the study of parthenocarpic fruit set. Numbers 1-5 represent each of the five individual greenhouses utilized during experiment 1. The numbers listed inside each shaded box indicate the 2A×Gy8 F_3 family or control line membership of each plant. Yellow blocks indicate plants that were either damaged or failed to reach maturity. Parental line '2A' is denoted as "P1", while parental line 'Gy8' is denoted as "NP1".



Addendum 5. Heat map displaying the spatial position and the number of initiated fruits set on each experimental plant during experiment 2 in the study of parthenocarpic fruit set. Numbers 1-5 represent each of the five individual greenhouses utilized during experiment 2. The numbers listed inside each shaded box indicate the 2A×Gy8 F_3 family or control line membership of each plant. Yellow blocks indicate plants that were either damaged or failed to reach maturity. Parental line '2A' is denoted as "P1", while parental line 'Gy8' is denoted as "NP1".



Addendum 6. Heat map displaying the spatial position and the number of initiated fruits set on each experimental plant during experiment 3 in the study of parthenocarpic fruit set. Numbers 1-5 represent each of the five individual greenhouses utilized during experiment 3. The numbers listed inside each shaded box indicate the 2A×Gy8 F_3 family or control line membership of each plant. Yellow blocks indicate plants that were either damaged or failed to reach maturity. Parental line '2A' is denoted as "P1", while parental line 'Gy8' is denoted as "NP1".



Addendum 7. The maturity of non-parthenocarpic cucumber cultivars at the time of harvest in 2012 commercial production trials at the University of Wisconsin-Madison Agricultural Research Station in Hancock, WI. Maturity was measured by the number of nodes on the main stem. Ten plants from each of three commercial plots were sampled. Two of the plots were planted with the same 'Excursion' cultivar.

Cultivar	Entry Number	Number of Nodes at Harvest
Vlaspik	1	18
	2	18
	3	18
	4	17
	5	18
	6	19
	7	19
	8	18
	9	18
	10	18
Excursion	1	19
	2	18
	3	21
	4	18
	5	24
	6	19
	7	18
	8	20
	9	20
	10	21
Excursion	1	18
	2	20
	3	19
	4	21
	5	20
	6	21
	7	20
	8	19
	9	22
	10	20
Overall Average		19.3

Addendum 8. 2A×Gy8 F_3 family means for each experiment conducted to study parthenocarpic fruit set in cucumber.

Family	Exp 1	Exp 2	Exp 3	Exp 2 and 3[z]	Pooled[y]
1	3.40	5.33	5.00	5.17	4.36
2	3.40	3.00	4.00	3.50	3.45
3	5.80	4.00	5.00	4.50	5.09
4	4.80	4.67	7.00	5.83	5.36
5	2.00	2.67	3.33	3.00	2.55
6	4.60	5.00	3.33	4.17	4.36
7	1.40	2.00	4.00	3.00	2.27
8	1.40	1.67	3.00	2.33	1.91
9	2.40	3.67	4.33	4.00	3.27
10	5.00	4.67	3.67	4.17	4.55
11	1.00	2.67	3.00	2.83	2.00
12	0.60	5.00	4.00	4.40	2.50
13	3.20	4.33	5.33	4.83	4.09
14	3.20	6.00	3.33	4.67	4.00
15	2.80	6.00	4.00	5.00	4.00
16	1.60	3.00	2.33	2.67	2.18
17	1.80	2.67	3.33	3.00	2.45
18	1.75	4.00	3.00	3.50	2.80
19	1.20	2.67	2.00	2.33	1.82
20	1.60	3.67	4.33	4.00	2.91
21	4.40	3.67	5.33	4.50	4.45
22	1.80	2.67	1.33	2.00	1.91
23	3.60	5.33	3.33	4.33	4.00
24	1.80	1.67	3.00	2.33	2.09
25	4.80	5.67	4.33	5.00	4.91
26	2.00	3.67	4.33	4.00	3.09
27	2.00	3.00	4.00	3.50	2.82
28	3.40	2.67	2.33	2.50	2.91
29	6.00	5.67	3.33	4.50	5.18
30	1.80	3.00	2.33	2.67	2.27
31	2.40	2.33	3.67	3.00	2.73
33	2.50	3.33	3.33	3.33	3.00
34	6.40	6.00	4.33	5.17	5.73
35	4.80	3.33	3.33	3.33	4.00
36	3.00	6.00	2.67	4.00	3.50
37	3.40	3.00	4.33	3.67	3.55

Family	Exp 1	Exp 2	Exp 3	Exp 2 and 3[z]	Pooled[y]
38	3.50	4.00	5.33	4.67	4.20
39	3.00	3.33	3.33	3.33	3.18
40	0.80	3.67	3.00	3.33	2.18
41	4.60	4.33	4.33	4.33	4.45
42	3.40	4.00	3.67	3.83	3.64
44	5.80	4.00	4.33	4.17	4.91
45	2.20	5.33	4.67	5.00	3.73
46	5.00	4.00	3.33	3.67	4.27
47	2.40	3.00	3.00	3.00	2.73
49	5.20	4.00	4.33	4.17	4.64
50	1.40	3.33	3.67	3.50	2.55
51	3.40	3.00	3.67	3.33	3.36
52	3.67	3.33	4.67	4.00	3.89
53	1.80	2.67	2.33	2.50	2.18
54	2.20	1.67	2.33	2.00	2.09
55	5.80	3.67	3.67	3.67	4.64
56	7.40	4.67	7.50	5.80	6.60
57	3.80	4.33	5.67	5.00	4.45
58	2.60	3.67	4.67	4.17	3.45
59	5.00	4.67	5.67	5.17	5.09
60	4.00	2.00	4.00	3.00	3.45
61	3.60	3.00	4.00	3.50	3.55
62	6.00	3.67	4.00	3.80	4.90
63	2.80	3.67	3.33	3.50	3.18
64	5.00	4.00	7.00	5.50	5.27
65	1.60	4.00	4.00	4.00	2.91
66	2.20	3.50	2.33	2.80	2.50
67	1.60	2.00	3.00	2.50	2.09
68	3.40	6.33	5.67	6.00	4.82
69	NA	2.67	2.33	2.50	2.50
72	2.60	3.00	2.00	2.50	2.55
73	2.00	3.33	2.00	2.67	2.36
74	5.60	4.00	6.67	5.33	5.45
75	5.80	3.67	3.00	3.33	4.45
76	1.80	3.67	2.67	3.17	2.55
77	0.80	1.67	3.00	2.33	1.80
78	2.60	5.00	5.00	5.00	3.91
79	1.00	2.33	3.33	2.83	2.10

Family	Exp 1	Exp 2	Exp 3	Exp 2 and 3[z]	Pooled[y]
80	2.00	4.67	2.00	3.33	2.73
81	3.20	3.33	5.00	4.17	3.73
82	2.75	3.67	4.00	3.83	3.40
83	0.80	3.00	2.00	2.50	1.73
84	4.20	5.33	3.33	4.33	4.27
85	3.20	3.67	5.00	4.33	3.82
86	4.00	3.50	4.67	4.20	4.10
87	2.20	3.67	3.67	3.67	3.00
88	2.00	5.00	4.67	4.83	3.55
89	1.40	2.67	2.00	2.33	1.91
90	4.00	5.00	5.33	5.17	4.70
91	4.75	3.33	4.67	4.00	4.30
92	1.60	3.00	3.33	3.17	2.45
93	4.00	6.67	5.00	5.83	5.00
94	5.60	4.33	7.00	5.67	5.64
95	1.40	2.67	3.00	2.83	2.18
97	3.80	5.33	5.67	5.50	4.73
98	3.00	3.67	2.67	3.17	3.09
99	3.40	4.33	3.67	4.00	3.73
100	3.20	3.00	4.00	3.50	3.36
101	2.60	4.33	2.00	3.17	2.91
102	3.20	3.33	4.67	4.00	3.64
103	3.80	4.33	3.67	4.00	3.91
104	3.20	4.33	3.00	3.67	3.45
105	2.75	2.67	2.67	2.67	2.70
107	5.40	4.67	4.00	4.33	4.82
108	4.40	1.67	3.00	2.33	3.27
109	4.60	3.33	4.00	3.67	4.09
110	4.20	4.33	2.67	3.50	3.82
111	1.80	2.00	3.33	2.67	2.27
113	4.20	6.00	4.33	5.17	4.73
114	1.75	2.33	4.00	3.17	2.60
115	3.00	2.67	2.33	2.50	2.73
116	1.40	3.00	2.67	2.83	2.18
117	0.80	3.67	4.67	4.17	2.64
118	6.20	4.33	5.67	5.00	5.55
119	2.00	5.33	2.33	3.83	3.00
120	3.60	3.33	2.67	3.00	3.27

Family	Exp 1	Exp 2	Exp 3	Exp 2 and 3[z]	Pooled[y]
121	3.40	4.00	3.00	3.50	3.45
122	4.80	4.67	5.67	5.17	5.00
123	4.80	2.67	2.33	2.50	3.55
124	0.80	3.67	3.33	3.50	2.27
125	3.20	6.67	4.67	5.67	4.55
126	6.40	1.67	3.67	2.67	4.36
127	0.60	1.33	2.00	1.67	1.18
128	7.40	6.00	4.67	5.33	6.27
129	2.00	4.33	4.67	4.50	3.36
131	2.00	1.67	1.67	1.67	1.82
133	3.60	4.00	3.67	3.83	3.73
134	4.20	4.00	4.00	4.00	4.09
135	3.40	4.33	4.00	4.17	3.82
136	1.00	1.00	1.67	1.40	1.20
137	3.60	5.67	3.00	4.33	4.00
138	7.00	6.33	5.00	5.67	6.00
139	4.00	3.00	3.00	3.00	3.45
140	6.25	6.67	3.33	5.00	5.50
141	2.60	2.33	2.67	2.50	2.55
142	2.40	3.00	3.33	3.17	2.82
143	3.80	2.67	2.33	2.50	3.09
144	1.60	3.00	3.67	3.33	2.55
145	0.80	2.67	2.33	2.50	1.73
146	3.60	2.33	2.67	2.50	3.00
147	2.00	4.67	4.67	4.67	3.45
148	2.80	4.67	5.33	5.00	4.00
149	2.80	4.00	4.33	4.17	3.55
150	2.80	4.67	3.00	3.83	3.36
151	3.00	2.00	3.33	2.67	2.82
152	5.20	3.33	2.67	3.00	4.00
153	2.00	3.33	2.33	2.83	2.50
154	5.80	5.33	4.33	4.83	5.27
155	NA	1.50	1.67	1.60	1.60
156	1.75	2.00	4.00	3.00	2.50
157	1.60	3.33	3.67	3.50	2.64
158	1.20	2.33	3.33	2.83	2.20
161	2.80	3.33	3.00	3.17	3.00
162	1.60	2.00	2.67	2.33	2.00

Family	Exp 1	Exp 2	Exp 3	Exp 2 and 3[z]	Pooled[y]
164	2.60	7.00	4.33	5.67	4.27
165	1.40	2.33	3.00	2.67	2.09
167	4.00	4.67	5.00	4.83	4.45
168	2.00	3.00	4.67	3.83	3.00
169	4.50	5.00	4.67	4.83	4.70
170	4.80	4.00	4.67	4.33	4.55
171	1.20	1.33	2.33	1.83	1.70
172	4.25	5.00	4.00	4.50	4.40
173	2.40	4.33	3.67	4.00	3.20
174	7.60	4.00	5.00	4.50	5.91
175	2.00	4.33	3.33	3.83	3.00
176	2.50	4.33	5.33	4.83	3.90
177	0.75	2.67	2.33	2.50	1.80
178	7.00	4.33	4.67	4.50	5.64
179	1.60	4.00	3.00	3.50	2.64
180	3.20	3.67	4.67	4.17	3.73
181	1.40	3.00	2.67	2.83	2.18
182	1.80	3.00	4.00	3.50	2.73
184	0.75	1.33	1.00	1.17	1.00
185	5.00	2.67	4.00	3.33	4.09
186	3.40	3.67	3.67	3.67	3.55
187	3.20	6.00	5.33	5.67	4.55
188	1.80	3.33	3.33	3.33	2.64
189	2.00	3.33	3.00	3.17	2.64
190	3.00	6.67	5.67	6.17	4.90
191	NA	1.33	1.67	1.50	1.50
192	2.00	2.33	3.00	2.67	2.36
193	3.20	3.00	4.00	3.50	3.36
196	3.00	4.00	5.67	4.83	4.00
198	5.00	4.33	4.67	4.50	4.73
199	7.20	4.67	3.67	4.17	5.55
200	6.20	4.33	4.67	4.50	5.27
201	6.20	3.00	4.33	3.67	4.82
202	7.20	3.67	5.33	4.50	5.73
203	2.25	2.33	3.00	2.67	2.50
205	NA	1.50	4.00	3.00	3.00
206	2.20	3.00	4.33	3.67	3.00
207	4.60	3.00	4.67	3.83	4.18

Family	Exp 1	Exp 2	Exp 3	Exp 2 and 3[z]	Pooled[y]
208	2.20	2.00	2.00	2.00	2.09
209	3.00	2.00	2.33	2.17	2.55
210	2.40	2.00	3.33	2.67	2.55
211	3.80	2.67	2.33	2.50	3.09
212	3.25	3.00	3.67	3.33	3.30
213	3.00	5.00	3.00	4.00	3.55
214	3.00	3.33	4.67	4.00	3.60
215	3.40	5.33	3.00	4.17	3.82
216	5.40	6.00	6.00	6.00	5.73
217	4.33	4.33	6.00	5.17	4.89
218	3.80	3.00	4.33	3.67	3.73
219	1.40	2.67	1.67	2.17	1.82
220	3.40	4.33	3.67	4.00	3.73
221	2.40	4.00	3.67	3.83	3.18
222	5.00	3.67	3.00	3.33	4.00
223	3.80	4.00	4.00	4.00	3.91
224	2.40	3.33	2.33	2.83	2.64
2A	3.78	6.50	5.95	6.23	5.52
Gy8	3.10	3.15	3.00	3.08	3.08
2A×Gy8 F1	2.75	2.80	3.20	3.00	2.93

[z]F_3 family means from experiments 2 and 3 combined.

[y]F_3 family means from experiments 1-3 pooled.

Chapter 2

Construction of a Linkage Map in an $F_{2:3}$ Population Segregating for Parthenocarpic Fruit Set in Cucumber (*Cucumis sativus* L.)

Abstract

The construction of linkage maps in cucumber for the identification of QTL and potential candidate genes for economically important traits has become more functional and prolific with the recent release of whole genome sequence data and the development of thousands of co-dominant SSR molecular markers. In this study we have developed a moderately saturated linkage map for use in identifying QTL associated parthenocarpic fruit set in cucumber. A mapping population consisting of 205 F_3 families was generated from the cross of a highly parthenocarpic inbred line, '2A', with a non-parthenocarpic inbred line, 'Gy8'. Despite the low level of polymorphism (6.65%) between the two parental lines, a linkage map consisting of 185 SSR, 5 STS, and 2 dCAPS marker loci in seven linkage groups covering 571.7 cM was developed. Measured in physical distance, the linkage map covered 164.3 Mb and accounted for approximately 85% of the distance covered by the assembled chromosomes in the Gy14 Draft Genome Assembly Version 1.0 (193.2 Mb). The linkage map has an average marker interval of 3 cM. In addition, with the recent publication by Sun et al. (2006b) utilizing an independent population developed from the same parental lines, comparisons could be made to validate the observed low levels of polymorphism and genomic regions lacking polymorphism in this study.

Introduction

Cucumber has a number of characteristics that make it ideal for genetic and marker assisted selection (MAS) studies. Cucumber has a short life cycle of approximately 90 days from seed to seed and is easy to grow. Furthermore, cultivated cucumber does not appear to suffer from inbreeding depression, although it is monoecious and an outcrossing species (Cramer and Wehner, 1999; Jenkins, 1942; Robinson and Decker-Walters, 1997; Rubino and Wehner, 1986). This characteristic is favorable for genetic studies and maintenance of genetic stocks. From a genomic perspective, cucumber has a relatively small genome size of 367 Mb with seven chromosome pairs ($2n = 2x = 14$) (Arumuganathan and Earle, 1991). In addition, there is evidence that cucumber has not had a recent genome duplication event and a majority of genes appear as single copy genes throughout the genome (Huang et al., 2009). Cucumber also benefits from a wealth of knowledge from previous studies, as the crop has served as a model species for studying plant biological processes such as sex determination, plant vascular physiology, and organellar genomics (Alverson et al., 2011; Havey, 1997; Havey et al., 1998; Lough and Lucas, 2006; Tanurdzic and Banks, 2004; Wang et al., 2010; Xoconostle-Cazares et al., 1999; Zhang et al., 2010a).

A major obstacle to previous genetic mapping efforts has been the narrow genetic base of cucumber. Past evaluations of genetic diversity in cucumber have reported low degrees of variation between 3 and 12% (Dijkuizen et al., 1996; Horejsi and Staub, 1999; Knerr et al., 1989; Meglic et al., 1996; Meglic and Staub, 1996). As expected, commercial varieties were found to have an extremely narrow genetic base in these studies. However, a recent study by Lv et al. (2012) utilized SSR markers to analyze a diverse mega collection of 3342 cucumber accessions

from various international germplasm collections and identified three distinct population groups (India, China/East Asia, and North America/Europe/West Asia) with higher estimates of overall diversity exceeding 20%. A large amount of variation was found not only within, but also between the population groups and suggests that crosses made outside of effective heterotic groups may provide new sources of variation. Interestingly, a high level of homogeneity within each population group was noted and may reflect past genetic bottlenecks and inbreeding within cucumber populations. These findings suggest that heterosis can be obtained if wide and diverse crosses are employed (Ghaderi and Lower, 1979a; Ghaderi and Lower, 1979b; Hayes and Jones, 1916; Hutchins, 1938; de Lalla et al., 2010; Singh et al., 2012).

The first linkage maps in cucumber were constructed using phenotypic data to link morphological traits (Fanourakis and Simon, 1987; Pierce and Wehner, 1990; Vokalounakis, 1992). These maps only consisted of a few simply inherited trait loci and were difficult to utilize in MAS due to small numbers of loci and weak linkages between traits. The development of molecular markers enabled the development of more saturated linkage maps and the identification of stronger linkages between traits and map loci. The first uses of molecular markers in linkage mapping utilized predominantly isozyme and restriction fragment length polymorphism (RFLP) markers (Kennard et al., 1994; Knerr and Staub, 1992; Meglic and Staub, 1996). These early maps provided a foundation for mapping economically important traits in cucumber, but the high costs and limited availability of these markers restricted researchers to construction of sparsely populated linkage maps. Soon linkage maps expanded to the use of random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) markers which were able to inexpensively generate multiple map loci per marker and did not require any prior knowledge of sequence. Using RAPD and AFLP markers, the first

moderately saturated linkage maps were produced for use in quantitative trait locus (QTL) identification and MAS (Fazio et al., 2003; Horejsi et al., 2000; Park et al., 2000; Serquen et al., 1997; Sun et al., 2006b).

In recent years, there has been rapid progress in the development of genetic and genomic resources in cucumber. The release of whole genome sequences for three cucumber lines, North American processing type 'Gy14', Northern China fresh market type '9930', and North European type 'B10', and the development of a large amount of co-dominant simple sequence repeat (SSR) molecular markers have made genetic mapping and gene cloning in cucumber much easier than before (Cavagnaro et al., 2010; Huang et al., 2009; Ren et al., 2009; Woycicki et al., 2011; Yang et al., 2012). With the introduction of sequencing technologies, RAPD and AFLP markers were largely replaced with SSR, single nucleotide polymorphism (SNP), and sequenced characterized amplified region (SCAR) markers because of their co-dominant nature and improved reproducibility across differing populations. By utilizing cucumber genome sequence data and large collections of inexpensive SSR markers, numerous studies have been able to construct linkage maps to identify QTL and in some cases candidate genes for horticulturally important traits (Amano et al., 2013; He et al., 2013; Kang et al., 2011; Li et al., 2011; Li et al., 2013; Miao et al., 2011; Zhang et al., 2010b; Zhang et al., 2013). The large number of SSR loci mapped by various studies and their easy transference across populations has led to the construction of highly saturated consensus linkage maps (Ren et al., 2009; Yang et al., 2013). These consensus maps can be used to overcome problems with low genetic diversity found in single cross populations by exploiting genetic diversity found in wide ranging populations. Saturated consensus maps used in combination with available whole genome

sequence data are providing a valuable resource for QTL identification, map-based gene cloning, association mapping, and MAS in cucumber.

For this study, SSR markers were chosen for construction of a linkage map for later use in identifying QTL associated with parthenocarpic fruit set in cucumber. SSR markers were chosen because they are readily available and are relatively inexpensive to utilize. The main objective during construction of the linkage map was to maximize genome coverage while placing marker loci at regular intervals throughout the map. In genomic regions where available polymorphic SSR markers were exhausted, sequence tag site (STS) and derived cleaved amplified polymorphic sequence (dCAPS) markers were synthesized to supplement the SSR based linkage map. A second objective of this study was to increase marker saturation in genomic regions identified as potential QTL for refinement of QTL locations.

Materials and Methods

Mapping Population

An $F_{2:3}$ mapping population was created for linkage map construction and QTL identification from a cross between the highly parthenocarpic processing cucumber inbred line, '2A', and the non-parthenocarpic processing cucumber inbred line, 'Gy8' (Chapter 1 Addendum 1). For construction of the linkage map, 205 F_2 plants derived from the self-pollination of a single 2A×Gy8 F_1 plant were genotyped. For QTL identification (Chapter 3) and a preliminary QTL analysis preformed to aid in the construction of the linkage map, each F_3 plant was scored for the number of ovaries initiating parthenocarpic fruit development (refer to Chapter 1). The mean

value obtained for each F_3 family was assigned as the phenotype of the F_2 plant from which it was derived.

Molecular Marker Analysis

Genomic DNA was isolated from unexpanded young leaves. Leaf samples were first lyophilized and then ground into fine powder with a high-throughput homogenizer (OPS Diagnostics, Lebanon, NJ). Genomic DNA was then extracted from the ground tissue with the CTAB method and purified with phenol/chloroform (Murray and Thompson, 1980). The DNA concentration of all samples was determined using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE). Samples were then diluted with 1X TE Buffer (pH 8.0) to a concentration of 25 ng μL^{-1}.

All polymerase chain reactions (PCR) were performed using an Applied Biosystems 2720 thermal cycler (Applied Biosystems, Foster City, CA). Each PCR reaction consisted of: 1 µL of diluted DNA (25ng/µL), 1µL of 1X PCR buffer (Fermentas, Glen Burnie, MD), 0.5µL each of 5 µM forward and reverse primers, 0.2µL of 10 µM dNTPs, 0.5U of Dream Taq *Taq* polymerase (5U/µL) (Fermentas, Glen Burnie, MD), and 6.7 µL of water for a final reaction volume of 10 µL. A touchdown PCR program detailed by Weng et al. (2005) was utilized for all primer sets. The program is as follows: 3 min initial denaturation at 95°C; six cycles of 45 s at 94°C for denaturation, 5 min at 68°C for annealing, 1 min at 72°C for extension, with the annealing temperature being reduced by 2°C per cycle; eight cycles of 45 s at 94°C for denaturation, 2 min at 58°C for annealing, 1 min at 72°C for extension, with the annealing temperature reduced by 1°C per cycle; a final 25 cycles of 45 s at 94°C for denaturation, 2 min at 50°C for annealing, and 1 min at 72°C for extension (Weng et al., 2005). PCR amplicons

obtained through dCAPS markers were digested with the appropriate restriction enzyme prior to gel electrophoresis. PCR amplicons from all primer sets were size-fractionated in denaturing polyacrylamide gels as described by Chen et al. (1998) with the exception of the use of 9% denaturing polyacrylamide gel prepared from stock solutions. Visualization of banding patterns was achieved by silver staining as described by Bassam et al. (1991) and modified by Weng and Lazar (2002). Banding patterns were scored manually and digital photographs were produced for long term preservation and reference (Addendum 1). Only two alleles were observed at each marker locus in this population.

Linkage analysis was conducted using JoinMap 3.0 software (Van Ooijen and Voorrips, 2001). Marker groups were calculated using the independence test logarithm of odds (LOD) with a minimum threshold of 4.0. Linkage groups and genetic distances were calculated by the regression mapping algorithm and the Kosambi mapping function with the following thresholds: linkage larger than a 1.0 LOD value, recombination frequency of 0.400, and a goodness of fit jump of 5.0 (Kosambi, 1943). A ripple function was performed after the addition of each marker locus to construct an optimized marker order.

Whole Genome Re-sequencing Data

The parental lines ‘2A’ and ‘Gy8’ were re-sequenced with the Illumina HiSeq 2000 platform (Illumina Inc., San Diego, CA) using the 2×100 base paired ends module with a mean coverage of 10×. The resulting short reads produced for ‘2A’ and ‘Gy8’ were mapped to the Gy14 Draft Genome Assembly Version 1.0 as a reference genome using Bowtie Short Read Alignment Software, Version 0.12.8 (Langmead et al., 2009; Yang et al., 2012). After alignment, Bowtie was used to identify indel and SNP polymorphisms between ‘2A’ and ‘Gy8’.

Molecular Markers for Linkage Analysis

For map construction, 3532 SSR markers previously developed from the genome sequences of cucumber inbred lines 'Gy14' and '9930' were selected for polymorphism screening (Cavagnaro et al., 2010; Ren et al., 2009). Linkage groups were matched to the corresponding chromosome using the Gy14 draft genome assembly in accordance to Yang et al. (2012). After a rough linkage map was assembled using the available polymorphic SSR markers, an additional 153 indel-derived STS markers and 21 dCAPS markers were designed from polymorphisms identified between '2A' and 'Gy8' with Bowtie from whole genome re-sequencing data (Michaels and Amisino, 1998). These additional markers were designed to fill large gaps identified in the rough linkage map and to increase marker density in genomic regions identified as preliminary QTL by an initial QTL mapping analysis performed using a preliminary linkage map. Selected indel polymorphisms were required to be a minimum of five base pairs (bp) in length and to have unanimous agreement among aligned sequence reads. Indel polymorphisms with higher sequence read coverage were preferentially selected and typical coverage for those selected was between 7-15×. SNP polymorphisms were held to the same selection criteria other than a length requirement. STS marker primers were designed with primer design software, Primer 3 (Rozen and Skaletsky, 2000). dCAPS marker primers containing a restriction enzyme cut site around the SNP polymorphism were designed using dCAPS Finder 2.0 software (Neff et al., 2002). Since dCAPS Finder 2.0 only identifies a single primer containing the polymorphism, Primer 3 was used to design the second primer of the primer set. Both STS and dCAPS primer sets were designed to amplify regions ranging in length from 140 to 250 bp. To verify the specificity of the newly designed STS and dCAPS markers, *in silico* PCR was performed to

confirm single PCR amplicons using the Gy14 and 9930 draft genome assemblies as templates (Huang et al., 2009; Yang et al., 2012).

Results and Discussion

General Assessment of the Linkage Map

The linkage map constructed by this study for the investigation of parthenocarpic fruit set in cucumber contains 192 marker loci consisting of 185 SSR, 5 STS, and 2 dCAPS markers contained within the expected seven linkage groups (Figure 1). The linkage map covers a total map length of 571.7 cM as calculated using the Kosambi mapping function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001). Measured in physical distance, the linkage map covers 164.3 Mb and accounts for approximately 85% of the distance covered by the assembled chromosomes in the Gy14 draft genome assembly (193.2 Mb) (Table 1). The linkage map has an average marker density of one marker locus every 3 cM (Table 1).

Of the 3532 SSR markers screened for polymorphisms, 235 (6.65%) were found to be polymorphic between the parental lines '2A' and 'Gy8' (Table 2). Since many of the polymorphic SSR markers were located in close proximity to each other, some identified polymorphic markers were excluded from linkage map construction because they were regarded as duplicate markers. However, in genomic regions identified as preliminary QTL by an initial QTL mapping analysis, all SSR markers were analyzed and included in the linkage map. In addition, due to the suspected poor quality and low coverage of the re-sequencing data, very few of the STS and dCAPS markers successfully identified true polymorphisms (7/174 or 4.0%). All STS and dCAPS markers identified as polymorphic were included in the linkage map.

Significant segregation distortion was observed for marker loci along a large portion of chromosome 5 (Table 3). Segregation distortion is a common phenomenon and can result from changes in fertility of either gametes or zygotes and may also be a consequence of environmental factors (Lyttle, 1991; Xu et al., 1997). A clear explanation for segregation distortion along chromosome 5 in this population is not immediately evident but it is interesting to note in combination with the two large physical intervals unaccounted for by the linkage map that lie on either side of the region showing distortion (Table 4, Table 3). Chromosomes 1 and 4 also contain a small number of marker loci showing segregation distortion (Table 3).

There are numerous intervals larger than 3 Mb without marker coverage in the linkage map (Table 4). The largest intervals unaccounted for by the linkage map occur on chromosomes 1 and 5. None of the chromosomes show strong evidence for the clustering of markers due to suppressed recombination (Table 1, Table 5). Any perceived clustering of markers on the linkage map is mostly explained by the inclusion of markers in close physical proximity of each other; due to the uneven distribution of available polymorphic markers. Table 4 presents the number of SSR, STS, and dCAPS markers screened for polymorphisms within each interval larger than 3 Mb that lacked marker coverage by the linkage map. No polymorphisms could be detected in these regions with the available markers. An attempt to minimize the size of these intervals by the addition of markers in flanking regions inadvertently led to numerous markers located in close physical proximity. In addition, each chromosome has comparable measurements of physical chromosome distance covered per centimorgan, suggesting the absence of major regions of recombination promotion or suppression (Table 1). Comparison of the marker loci order and relative positions in the 2A×Gy8 linkage map with marker loci

positions in the Gy14 draft genome assembly also support the observation of normal recombination frequencies and absence of major marker clustering (Table 5).

General Assessment of Large Intervals Unaccounted for by the Linkage Map

A comparison of the Gy14 draft genome assembly and the linkage map shows that the linkage map contains a number of relatively large intervals lacking marker coverage (Table 4, Table 5). While it is possible that polymorphisms do exist in these regions, the number of loci screened in these regions without the successful identification of a polymorphism is well below the average expected polymorphism rate of 6.65% for this population (Table 2). Common ancestry between the two parental lines, '2A' and 'Gy8', could be potentially contributing to genomic regions of low polymorphism in this population. Although a direct common ancestor could not be identified, both inbred lines have pedigrees including numerous lines developed thru the cucumber research programs at the University of Wisconsin-Madison (Chapter 1 Addendum 1). Both parental lines also include close relationships with the public gynoecious inbred line series (e.g. the Gy series), with '2A' descending from 'Gy7' and 'Gy8' itself descending from 'Gy14' (Chapter 1 Addendum 1). In addition, breeding efforts to incorporate traits relating to fruit quality, disease resistance, and favorable processing characteristics into elite processing cucumber lines may have also led to the incorporation of genomic regions derived from similar sources (Chapter 1 Addendum 1).

The 2A×Gy8 linkage map constructed here will ultimately be utilized for identification of QTL associated with parthenocarpic fruit set in cucumber. The large intervals in the linkage map without marker coverage may hinder the ability to detect QTL. This is especially concerning for the detection of minor QTL. However, the presence of any major QTL will likely be detectable

even with large unaccounted for intervals on the linkage map. The inheritance of parthenocarpic fruit set in cucumber is unclear as past studies have suggested models ranging from single gene inheritance to complex multigenic inheritance; with a recent QTL study by Sun et al. (2006b) indicating complex inheritance with numerous small QTL (El-Shawaf and Baker, 1981; Hawthorn and Wellington, 1930; Kvasnikov et al., 1970; Meshcherov and Juldasheva, 1974; Pike and Peterson, 1969; de Ponti and Garretson, 1976; Sun et al., 2006a; Sun et al., 2006b). While Sun et al. (2006b) utilized another 2A×Gy8 $F_{2:3}$ population derived from the same parental lines as the population used here, it should be noted that the current study employed a new approach to phenotypic evaluation of parthenocarpic fruit set in an attempt to limit confounding traits related to environment and yield that were significant in the previous study. Therefore, there is not any expectation for the presence of major or minor QTL in this study. Ultimately, the expectation is that QTL will be identified in the regions of the linkage map showing polymorphisms between the two parental lines.

Comparison to Other *C. sativus* Linkage Maps

The 2A×Gy8 linkage map presented here is similar in the number of marker loci and map distance covered to other recent linkage maps constructed for F_2 and RIL cucumber populations (Amano et al., 2013; He et al., 2013; Kang et al., 2011; Li et al., 2011; Miao et al., 2011; Weng et al., 2010; Zhang et al., 2010b). The level of polymorphism observed (6.65%) is relatively low when compared to other published single cross mapping populations (6.4%-17.0%) (Fazio et al., 2003; He et al., 2013; Kennard et al. 1994; Li et al., 2011; Miao et al., 2011; Serquen et al., 1997; Sun et al., 2006b; Weng et al., 2010; Zhang et al., 2010b) (Table 2). In agreement with this study, Sun et al. (2006b) observed a polymorphism rate of 6.68% for predominantly AFLP

and RAPD markers in another 2A×Gy8 $F_{2:3}$ population developed for the study of parthenocarpy. The lower genetic diversity seen in both of the 2A×Gy8 populations is likely due to the noted similarities in development of the parental lines, but may also reflect a suspected overall narrow genetic base among the majority of elite commercial processing type cucumber lines. The lack of polymorphism in this population may hinder future attempts at fine mapping QTL regions.

The large intervals lacking marker coverage in the 2A×Gy8 linkage map are similar to the intervals unaccounted for by the linkage map constructed by Sun et al. (2006b) (Addendum 2). The study by Sun et al. also observed few available polymorphic marker loci for chromosomes 1 and 5 (Table 1, Addendum 2). Correspondingly, the large unaccounted intervals 1.1, 1.2, 2.1, and 5.1 are also unaccounted for by the Sun et al. linkage map (Figure 1, Table 4, Addendum 2). These observations collaborate to add confidence to the accuracy of the 2A×Gy8 linkage map constructed by this study.

Future Focus

This study presents a moderately saturated linkage map with good utility for the identification of QTL contributing to parthenocarpic fruit set in cucumber. The quality of the map will ultimately be decided by the ability to successfully detect QTL accounting for a majority of the phenotypic variation observed for parthenocarpic fruit set. If sufficient QTL are not identified, one possible remedy includes screening more SSR, indel, and SNP loci in the large genomic regions lacking marker coverage in the current 2A×Gy8 linkage map, as numerous unexplored loci still exist. However, assuming the successful discovery of QTL for parthenocarpic fruit set, this linkage map will serve as a strong foundation for further fine mapping in the target genomic regions.

Literature Cited

Alverson, A.J., D.W. Rice, S. Dickinson, K. Barry, and J.D. Palmer. 2011. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. Plant Cell. 23:2499-2513.

Amano, M., A. Mochizuki, Y. Kawagoe, K. Iwahori, K. Niwa, J. Svoboda, T. Maeda, and Y. Imura. 2013. High-resolution mapping of *zym*, a recessive gene for *Zucchini yellow mosaic virus* resistance in cucumber. Theor. Appl. Genet. 126:2983-2993.

Arumuganathan, K. and E. Earle. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. 9:208-218.

Bassam, B. J., G. Caetano-Anolles, and P. M. Gresshoff. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal. Biochem. 196:80-83.

Chen, X. M., R. F. Line, and H. Leung. 1998. Genome scanning for resistance-gene analogs in rice, barley, and wheat by high-resolution electrophoresis. Theor. Appl. Genet. 97:345-355.

Cramer, C.S. and T.C. Wehner. 1999. Little heterosis for yield and yield components in hybrids of six cucumber inbreds. Euphytica. 110:99-108.

Cavagnaro, P.F., D.A. Senalik, L.M. Yang, P.W. Simon, T.T. Harkins, C.D. Kodira, S.W. Huang, and Y. Weng. 2010. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). BMC Genomics. 11:569.

Dijkhuizen, A., W.C. Kennard, M.J. Havey, and J.E. Staub. 1996. RFLP variability and genetic relationships in cultivated cucumber. Euphytica. 90:79-89.

El-Shawaf, I.I.S. and L.R. Baker. 1981. Inheritance of parthenocarpic yield in gynoecious pickling cucumber for once-over mechanical harvest by diallel analysis of six gynoecious lines. J. Amer. Soc. Hort. Sci. 106:365-370.

Fanourakis, N.E. and P.W. Simon. 1987. Analysis of genetic linkage in cucumber. J. Hered. 78:238-242.

Fazio, G., J.E. Staub, and M.R. Stevens. 2003. Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor. Appl. Genet. 107:864-874.

Ghaderi, A. and R.L. Lower. 1979a. Heterosis and inbreeding depression for yield in populations derived from six crosses of cucumber. J. Amer. Soc. Hort. Sci. 104:564-567.

Ghaderi, A. and R.L. Lower. 1979b. Analysis of generation means for yield in six crosses of cucumber. J. Amer. Soc. Hort. Sci. 104:567-572.

Havey, M.J. 1997. Predominant paternal transmission of the mitochondrial genome in cucumber. J. Hered. 88:232-235.

Havey, M.J., J.D. McCreight, B. Rhodes, and G. Taurick. 1998. Differential transmission of the *Cucumis* organellar genomes. Theor. Appl. Genet. 97:122-128.

Hawthorn, L.R. and R. Wellington. 1930. Geneva, a greenhouse cucumber that develops fruit without pollination. NY (Geneva) Agric Exp. Stat. Bull. 580:223-234.

Hayes, H.K. and D. F. Jones. 1916. First generation crosses in cucumbers. Connecticut Agric. Exp. Sta. Ann. Rep. 319-322.

He, X.M., Y.H. Li, S. Pandey, B.S. Yandell, M. Pathak, and Y. Weng. 2013. QTL mapping of powdery mildew resistance in WI 2757 cucumber (*Cucumis sativus* L.). Theor. Appl. Genet. 126:2149-2161.

Horejsi, T. and J.E. Staub. 1999. Genetic variation in cucumber (*Cucumis sativus* L.) as assessed by random amplified polymorphic DNA. Genet. Res. Crop Evol. 46:337-350.

Horejsi, T., J.E. Staub, and C. Thomas. 2000. Linkage of random amplified polymorphic DNA markers to downy mildew resistance in cucumber (*Cucumis sativus* L.). Euphytica. 115:105-113.

Huang, S., R. Li, Z. Zhang, and et al. 2009. The genome of the cucumber, *Cucumis sativus* L. Nat. Genet. 41:1275-1281.

Hutchins, A. E. 1938. Some examples of heterosis in cucumber, *Cucumis sativus* L. Proc. Amer. Soc. Hort. Sci. 36:660-664.

Jenkins, J.M., Jr., 1942. Natural self-pollination in cucumber. Proc. Amer. Soc. Hort. Sci. 40:411-412.

Kang, H.X., Y. Weng, Y.H. Yang, Z.H. Zhang, S.P. Zhang, Z.C. Mao, G.H. Cheng, X.F. Gu, S.W. Huang, and B.Y. Xie. 2011. Fine genetic mapping localizes cucumber scab resistance gene *Ccu* into an R gene cluster. Theor. Appl. Genet. 122:795-803.

Kennard, W.C., K. Poetter, A. Dijkhuizen, V. Meglic, J.E. Staub, and M.J. Havey. 1994. Linkages among RFLP, RAPD, isozyme, disease-resistance, and morphological markers in narrow and wide crosses of cucumber. Theor. Appl. Genet. 89:42-48.

Knerr, L.D., J.E. Staub, D.J. Holder, and B.P. May. 1989. Genetic diversity in *Cucumis sativus* L. assessed by variation at 18 allozyme coding loci. Theor. Appl. Genet. 78:119-128.

Kosambi, D. 1943. Estimation of map distances from recombination values. Ann. Eugen. 12:172-175.

Kvasnikov, B.V., N.T. Rogova, S.I. Tarakonova, and I. Ignatova. 1970. Methods of breeding vegetable crops under the covered ground. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 42:45-57.

de Lalla, J.G., V.A. Laura, S. Seabra, and A.I.I. Cardoso. 2010. Combining ability and heterosis of lines from Japanese cucumber type for yield. Hortic. Bras. 28:337-343.

Langmead, B., C. Trapnell, M. Pop, S.L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.

Li, Y.H., L.M. Yang, M. Pathak, D.W. Li, X.M. He, and Y. Weng. 2011. Fine genetic mapping of *cp*, a recessive gene for compact (dwarf) plant architecture in cucumber, *Cucumis sativus* L. Theor. Appl. Genet. 123:973-983.

Li, Y.H., C.L. Wen, and Y. Weng. 2013. Fine mapping of the pleiotropic locus *B* for black spine and orange mature fruit color in cucumber identifies a 50kb region containing a R2R3-MYB transcription factor. Theor. Appl. Genet. 126:2187-2196.

Lough, T.J. and W.J. Lucas. 2006. Integrative plant biology: role of phloem long-distance macromolecular trafficking. Annu. Rev. Plant Biol. 57:203-232.

Lower, R.L. 1996. Release of three gynoecious pickling cucumber inbreds: Gy7, Gy8, and Gy9. Cucurbit Genet. Coop. 19:94.

Lv, J., J. Qi, Q. Shi, D. Shen, S. Zhang, G. Shao, H. Li, Z. Sun, Y. Weng, Y. Shang, X. Gu, X. Li, X. Zhu, J. Zhang, R. van Treuren, W. van Dooijeweert, Z. Zhang, and S. Huang. 2012. Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). PLoS One. 7:e46919.

Lyttle, T.W. 1991. Segregation distorters. Annu. Rev. Genet. 25:511-557.

Meglic, V., F. Serquen, and J.E. Staub. 1996. Genetic diversity in cucumber (*Cucumis sativus* L.): I. A reevaluation of the U.S. germplasm collection. Genet. Res. Crop Evol. 43:533-546.

Meglic, V. and J.E. Staub. 1996. Genetic diversity in cucumber (*Cucumis sativus* L.): 2. An evaluation of selected cultivars released between 1846 and 1978. Genet. Res. Crop Evol. 43:547-558.

Meshcherov, E.T. and L.W. Juldasheva. 1974. Parthenocarpy in cucumber. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 51:204-213.

Miao, H., S.P. Zhang, X.W. Wang, Z.H. Zhang, M. Li, S.Q. Mu, Z.C. Cheng, R.W. Zhang, S.W. Huang, B.Y. Xie, Z.Y. Fang, Z.X. Zhang, Y. Weng, and X.F. Gu. 2011. A linkage map of cultivated cucumber (*Cucumis sativus* L.) with 248 microsatellite marker loci and seven genes for horticulturally important traits. Euphytica. 172:167-176.

Michaels, S.D. and R.M. Amisino. 1998. A robust method for the detecting single-nucleotide changes as polymorphic markers by PCR. Plant J. 14:381-385.

Murray, M.G. and W.F. Thompson. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8:4321-4325.

Neff, M.M., E. Turk, M. Kalishman. 2002. Web-based Primer Design for Single Nucleotide Polymorphism Analysis. Trends Genet. 18:613-615.

Park, Y.H., S. Senory, C. Wye, R. Antonise, J. Peleman, and M.J. Havey. 2000. A genetic map of cucumber composed of RAPDs, RFLPs, AFLPs, and loci conditioning resistance to papaya ringspot and zucchini yellow mosaic viruses. Genome. 43:1003-1010.

Pierce, L.K. and T.C. Wehner. 1990. Review of genes and linkage groups in cucumber. HortScience. 25:605-615.

Pike, L.M. and C.E. Peterson. 1969. Inheritance of parthenocarpy in the cucumber (*Cucumis sativus* L.) Euphytica. 18:101-105.

de Ponti, O.M.B. and F. Garretsen. 1976. Inheritance of parthenocarpy in pickling cucumbers (*Cucumis sativus* L.) and linkage with other characters. Euphytica. 25:633-642.

Ren, Y., Z.H. Zhang, J.H. Liu, J.E. Staub, Y.H. Han, Z.C. Cheng, X.F. Li, J.Y. Lu, H. Miao, H.X. Kang, B.Y. Bie, X.F. Gu, X.W. Wang, Y.C. Du, W.W. Jin, and S.W. Huang. 2009. An integrated genetic and cytogenetic map of the cucumber genome. PLoS One. 4:5795.

Robinson, R.W. and D.S. Decker-Walters. 1997. Cucurbits. CAB International, Wallingford, Oxon, UK.

Rozen, S. and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132:365-386.

Rubino, D. and T.C. Wehner. 1986. Effect of inbreeding on horticultural performance of lines developed from an open-pollinated pickling cucumber population. Euphytica. 35:459-464.

Serquen, F.C., J. Bacher, and J.E. Staub. 1997. Mapping and QTL analysis of a narrow cross in cucumber (*Cucumis sativus* L.) using random amplified polymorphic DNA markers. Mol. Breeding. 3:257-1365.

Singh, R., A. Singh, S. Kumar, and B.K. Singh. 2012. Heterosis and inbreeding depression for fruit characteristics in cucumber. Indian J. Hortic. 69:200-204.

Sun, Z. 2004. Inheritance and molecular marker-based genetic mapping of parthenocarpy in cucumber (*Cucumis sativus* L.). PhD. Thesis, University of Wisconsin-Madison.

Sun, Z., R.L. Lower, and J.E. Staub. 2006a. Variance component analysis of parthenocarpy in elite U.S. processing type cucumber (*Cucumis sativus* L.) lines. Euphytica. 148:331-339.

Sun, Z., J.E. Staub, S.M. Chung, and R.L. Lower. 2006b. Identification and comparative analysis of quantitative trait loci (QTL) associated with parthenocarpy in processing cucumber. Plant Breeding. 125:281-287.

Tanurdzic, M. and J.A. Banks. 2004. Sex-determining mechanisms in land plants. Plant Cell. 16 (Suppl.) S61-S71.

Van Ooijen, J.W. and R.E. Voorrips. 2001. JoinMap 3.0: Software for the calculation of genetic maps. Plant Research International, Wageningen, The Netherlands.

Vokalounakis, D.J. 1992. Heart leaf, a recessive leaf shape marker in cucumber: linkage with disease resistance and other traits. J. Hered. 83:217-221.

Wang, D.H., F. Li, Q.H. Duan, T. Han, Z.H. Xu, and S.N. Bai. 2010. Ethylene perception is involved in female cucumber flower development. Plant J. 61:862-872.

Weng, Y. and M.D. Lazar 2002. Amplified fragment length polymorphism- and simple sequence repeat-based molecular tagging and mapping of greenbug resistance gene *Gb3* in wheat. Plant Breeding. 121:218-223.

Weng, Y., W. Li, R.N. Devkota, and J.C. Rudd. 2005. Microsatellite markers associated with two *Aegilops tauschii*-derived greenbug resistance loci in wheat. Theor. Appl. Genet. 110:462-469.

Weng, Y., S. Johnson, J.E. Staub, and S.W. Huang. 2010. An extended microsatellite genetic map of cucumber, *Cucumis sativus* L. HortScience. 45:880-886.

Woycicki R, J. Witkowicz, P. Gawronski, J. Dabrowska, A. Lomsadze, M. Pawelkowicz, E. Siedlecka, K. Yagi, W. Plader, A. Seroczynska and et al. 2011. The genome sequence of the North-European cucumber (*Cucumis sativus* L.) unravels evolutionary adaptation mechanisms in plants. PLoS One. 6:e22728.

Xoconostle-Cazares, B., X. Yu, R. Ruiz-Medrano, H.L. Wang, J. Monzer, B.C. Yoo, K.C. McFarland, V.R. Franceschi, and W.J. Lucas. 1999. Plant paralog to viral movement protein that potentiates transport of mRNA into the phloem. Science. 283:94-98.

Xu, Y., L. Zhu, J. Xiao, N. Huang, and S.R. McCouch. 1997. Chromosomal regions associated with segregation distortion of molecular markers in F_2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Mol. Gen. Genet. 253:535-545.

Yang, L.M., D.H. Koo, Y.H. Li, X.J. Zhang, F.S. Luan, M.J. Havey, J.M. Jiang, and Y. Weng. 2012. Chromosome rearrangements during domestication of cucumber as revealed by high-density genetic mapping and draft genome assembly. Plant J. 71:895-906.

Yang, L.M., D.W. Li, Y.H. Li, X.F. Gu, S.W. Huang, J. Garcia-Mas, and Y. Weng. 2013. A 1,681-locus consensus genetic map of cultivated cucumber including 67 NB-LRR resistance gene homolog and ten gene loci. BMC Plant Biol. 13:53.

Zhang, B., V. Tolstikov, C. Turnbull, L.M. Hicks, and O. Fiehn. 2010a. Divergent metabolome and proteome suggest functional independence of dual phloem transport systems in cucurbits. Proc. Natl. Acad. Sci. 107:13532-13537.

Zhang, S.P., H. Miao, X.F. Gu, Y.H. Yang, B.Y. Xie, X.W. Wang, S.W. Huang, Y.C. Du, R.F. Sun, and T.C. Wehner. 2010b. Genetic mapping of the scab resistance gene in cucumber. J. Amer. Soc. Hort. Sci. 135:53-58.

Zhang, S.P., H. Miao, R.F. Sun, X.W. Wang, S.W. Huang, T.C. Wehner, and X.F. Gu. 2013. Localization of a new gene for bitterness in cucumber. J. Hered. 104:134-139.

Table 1. General statistics obtained for the linkage map generated for a 2A×Gy8 $F_{2:3}$ population developed for the study of parthenocarpic fruit set in *C. sativus*. The linkage map was constructed with 205 2A×Gy8 F_2 individuals and consists of 185 SSR, 5 STS, and 2 dCAPS marker loci in seven linkage groups covering 571.7 cM. Linkage groups were assigned to chromosomes according to Yang et al. (2012). Map distances were calculated using Kosambi's map function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001).

Linkage Group	Number of Marker Loci[z]	Length (cM)	Density (cM/Marker)	Physical Distance Covered (Mb)	Kb/cM	Total Length of CHR (Mb)[y]	Physical Distance Unaccounted For (Mb)
Chr1	15	82.0	5.47	23.52	286.9	28.50	4.97
Chr2	28	91.5	3.27	17.50	191.3	23.53	6.03
Chr3	31	108.8	3.51	39.47	362.8	40.31	0.84
Chr4	32	87.7	2.74	20.53	234.1	23.38	2.84
Chr5	18	54.7	3.04	20.42	373.4	27.52	7.10
Chr6	47	100.5	2.14	29.75	296.0	30.61	0.86
Chr7	21	46.5	2.21	13.08	281.4	19.34	6.25
Total	192	571.7	2.98	164.29	287.4	193.18	28.89

[z]Marker loci consist of predominantly SSR molecular markers accompanied by STS and dCAPS molecular markers.

[y]Total length of each chromosome reflects the total length of each chromosome in the Gy14 Draft Genome Assembly Version 1.0 (Yang et al., 2012).

Table 2. The number of SSR primer pairs showing polymorphisms between the parthenocarpic parent inbred line '2A' and the non-parthenocarpic parent inbred line 'Gy8' in 9% denaturing polyacrylamide gel electrophoresis (PAGE).

Primer Pairs Screened	Polymorphic Pairs[z]	Non-polymorphic Pairs	Overall Level of Polymorphism[y]
3532	235	3297	6.65%

[z]Primer pairs were scored as polymorphic if clear polymorphism was visible after two hours of fractionation in 9% denaturing polyacrylamide gel.

[y]Overall level of polymorphism refers to the level of polymorphism observed for the subset of primer pairs screened in this study.

Table 3. Table presenting the segregation data for each marker locus contained in the 2AxGy8 linkage map. The linkage map consists of 185 SSR, 5 STS, and 2 dCAPs marker loci in seven linkage groups. The "A" genotype has been assigned to '2A' while the "B" genotype represents 'Gy8'. Distortion from the expected 1:2:1 segregation for co-dominant markers is evaluated by the chi square test. The linkage map contains two SSR markers on Chromosome 6, UW084474 and UW026722, that could only be visually scored as dominant markers with the protocol outlined by this study. These markers were evaluated with an expected segregation ratio of 3:1.

CHR	Locus	Position[z]	A/A	A/B	B/B	A/- [y]	No Data	χ^2	Df	Sig.[x]
1	UW085383	0.0	48	104	49	0	4	0.25	2	-
1	SSR13109	1.0	46	108	50	0	1	0.86	2	-
1	SSR15108	1.7	44	110	51	0	0	1.58	2	-
1	SSR04644	4.3	44	102	53	0	6	0.94	2	-
1	SSR04304	4.9	46	107	50	0	2	0.75	2	-
1	SSR11654	9.0	50	98	54	0	3	0.34	2	-
1	SSR05793	10.6	47	101	49	0	8	0.17	2	-
1	UW045607	32.7	41	102	56	0	6	2.39	2	-
1	SSR15755	37.3	45	97	61	0	2	2.92	2	-
1	UW084360	39.7	46	96	51	0	12	0.26	2	-
1	UW083897	40.3	43	101	59	0	2	2.53	2	-
1	SSR14526	40.7	47	93	59	0	6	2.30	2	-
1	UW084542	41.4	48	99	57	0	1	0.97	2	-
1	UW083821	44.6	56	81	56	0	12	4.98	2	*
1	UW074644	82.0	49	101	54	0	1	0.26	2	-
2	UW084907	0.0	54	102	45	0	4	0.85	2	-
2	SSR00204	0.9	54	100	46	0	5	0.64	2	-
2	SSR18937	3.7	52	98	46	0	9	0.37	2	-
2	SSR13532	4.6	52	101	46	0	6	0.41	2	-
2	UW059395	7.2	58	105	42	0	0	2.62	2	-
2	UW043178	7.5	53	104	38	0	10	3.17	2	-
2	UW043203	7.6	51	92	36	0	26	2.65	2	-
2	UW085388	9.0	57	100	38	0	10	3.83	2	-
2	UW043299	9.1	58	104	41	0	2	2.97	2	-
2	UW084463	9.4	58	105	38	0	4	4.38	2	-
2	UWSTS0384	9.8	58	106	39	0	2	3.96	2	-
2	UW057528	10.6	59	105	39	0	2	4.18	2	-
2	SSR04869	11.1	57	108	40	0	0	3.41	2	-
2	UW085357	12.2	61	104	40	0	0	4.35	2	-
2	UW085360	13.4	53	98	39	0	15	2.25	2	-
2	SSR04870	16.4	57	102	43	0	3	1.96	2	-
2	UW078361	31.6	53	104	47	0	1	0.43	2	-
2	UW078335	32.0	53	106	46	0	0	0.72	2	-

CHR	Locus	Position[z]	A/A	A/B	B/B	A/- [y]	No Data	χ^2	Df	Sig.[x]
2	UW078088	33.0	50	105	42	0	8	1.51	2	-
2	UW053502	35.5	51	105	43	0	6	1.25	2	-
2	UW082700	41.3	46	107	45	0	7	1.30	2	-
2	SSR16916	49.9	44	100	42	0	19	1.10	2	-
2	UW036707	52.1	49	109	45	0	2	1.27	2	-
2	UW083968	52.4	46	107	46	0	6	1.13	2	-
2	UW016354	72.4	57	94	51	0	3	1.33	2	-
2	UW012751	83.6	52	91	48	0	14	0.59	2	-
2	SSR16028	88.9	51	98	52	0	4	0.13	2	-
2	SSR03606	91.5	47	92	47	0	19	0.02	2	-
3	SSR14159	0.0	51	104	47	0	3	0.34	2	-
3	SSR05312	5.8	51	107	44	0	3	1.20	2	-
3	SSR02451	6.5	51	107	40	0	7	2.52	2	-
3	UW085290	11.5	50	107	39	0	9	2.89	2	-
3	SSR16408	15.4	55	106	42	0	2	2.06	2	-
3	SSR05891	21.9	44	86	41	0	34	0.11	2	-
3	SSR14725	23.8	44	109	46	0	6	1.85	2	-
3	SSR01573	25.3	47	111	45	0	2	1.82	2	-
3	UW055751	25.9	48	109	46	0	2	1.15	2	-
3	SSR03409	28.3	42	109	45	0	9	2.56	2	-
3	UW083723	47.8	50	103	51	0	1	0.03	2	-
3	UW084166	47.8	50	103	48	0	4	0.16	2	-
3	SSR07220	48.4	49	102	50	0	4	0.05	2	-
3	SSR00525	53.0	50	106	47	0	2	0.49	2	-
3	SSR02068	53.1	49	108	48	0	0	0.60	2	-
3	SSR06210	54.8	47	113	45	0	0	2.19	2	-
3	UW083972	55.6	44	112	49	0	0	2.00	2	-
3	SSR16056	56.9	44	111	46	0	4	2.23	2	-
3	SSR02132	57.0	44	110	50	0	1	1.61	2	-
3	UW084363	61.0	47	102	55	0	1	0.63	2	-
3	UW083944	62.5	43	102	53	0	7	1.19	2	-
3	UW085394	76.2	45	98	45	0	17	0.34	2	-
3	SSR13949	79.6	52	94	50	0	9	0.37	2	-
3	SSR05328	83.1	46	101	46	0	12	0.42	2	-
3	SSR11397	83.2	48	99	44	0	14	0.42	2	-
3	SSR30236	100.2	55	105	45	0	0	1.10	2	-
3	SSR23159	102.5	52	107	46	0	0	0.75	2	-
3	SSR10783	106.0	53	108	44	0	0	1.38	2	-
3	SSR15312	106.5	51	103	44	0	7	0.82	2	-
3	SSR20578	106.8	52	106	43	0	4	1.41	2	-

CHR	Locus	Position[z]	A/A	A/B	B/B	A/- [y]	No Data	χ^2	Df	Sig.[x]
3	UW084555	108.8	47	109	46	0	3	1.28	2	-
4	SSR11074	0.0	48	101	53	0	3	0.25	2	-
4	UW083992	3.3	45	104	49	0	7	0.67	2	-
4	SSR05783	5.5	41	106	52	0	6	2.07	2	-
4	UW083734	6.4	42	105	58	0	0	2.62	2	-
4	UW084487	10.4	41	96	61	0	7	4.22	2	-
4	SSR05899	11.5	44	98	63	0	0	3.92	2	-
4	SSR01615	15.7	41	90	58	0	16	3.49	2	-
4	UW084453	22.0	44	94	67	0	0	6.57	2	**
4	SSR12386	30.2	39	96	61	0	9	5.02	2	*
4	UW084379	33.2	43	97	63	0	2	4.34	2	-
4	UW083899	34.6	42	96	66	0	1	6.35	2	**
4	SSR05415	38.1	40	92	59	0	14	4.04	2	-
4	SSR04482	41.5	44	93	55	0	13	1.45	2	-
4	UW029413	42.5	43	100	59	0	3	2.55	2	-
4	SSR13021	45.6	53	98	52	0	2	0.25	2	-
4	SSR04649	50.0	38	109	51	0	7	3.73	2	-
4	UW084520	50.5	39	114	51	0	1	4.24	2	-
4	SSR02697	50.6	38	110	51	0	6	3.91	2	-
4	UW083971	58.5	48	99	52	0	6	0.17	2	-
4	SSR14393	58.7	50	100	52	0	3	0.06	2	-
4	SSR10368	59.7	48	104	53	0	0	0.29	2	-
4	UW083893	69.2	56	81	40	0	28	4.16	2	-
4	UW083894	71.4	58	90	55	0	2	2.69	2	-
4	SSR05515	73.8	53	90	51	0	11	1.05	2	-
4	UW084851	81.2	58	94	51	0	2	1.59	2	-
4	SSR16498	81.9	54	92	52	0	7	1.03	2	-
4	SSR00249	82.8	56	95	53	0	1	1.05	2	-
4	SSR18551	83.2	58	93	53	0	1	1.83	2	-
4	SSR14054	84.7	58	86	53	0	8	3.43	2	-
4	SSR18559	86.0	58	96	50	0	1	1.33	2	-
4	UW084518	86.9	56	93	50	0	6	1.21	2	-
4	UW084519	87.7	58	94	51	0	2	1.59	2	-
5	UW084492	0.0	57	106	31	0	11	8.64	2	**
5	UW085421	1.3	57	103	33	0	12	6.84	2	**
5	UW084451	2.7	58	106	35	0	6	6.17	2	**
5	SSR11264	13.9	53	115	33	0	4	8.16	2	**
5	SSR16110	14.9	55	118	32	0	0	9.85	2	***
5	UW084566	15.2	53	113	32	0	7	8.41	2	**

CHR	Locus	Position[z]	A/A	A/B	B/B	A/- [y]	No Data	χ^2	Df	Sig.[x]
5	SSR32717	15.7	54	117	34	0	0	8.00	2	**
5	UW005172	16.8	45	107	33	0	20	6.10	2	**
5	SSR07711	17.4	56	106	32	0	11	7.61	2	**
5	SSR15321	20.4	50	115	38	0	2	5.01	2	*
5	SSR00182	25.7	49	118	36	0	2	7.03	2	**
5	UW001903	26.4	46	119	35	0	5	8.43	2	**
5	UW059902	32.3	46	115	40	0	4	4.54	2	-
5	SSR13409	52.6	59	93	49	0	4	2.11	2	-
5	SSR02895	52.9	57	95	50	0	3	1.20	2	-
5	UW084644	53.4	58	95	48	0	4	1.60	2	-
5	SSR19343	53.8	60	98	46	0	1	2.24	2	-
5	SSR20897	54.7	57	95	51	0	2	1.19	2	-
6	SSR16163	0.0	50	103	49	0	3	0.09	2	-
6	SSR02021	2.2	44	92	40	0	29	0.55	2	-
6	UWSTS0316	3.3	53	103	47	0	2	0.40	2	-
6	UW084474	9.3	0	0	50	152	3	0.00	1	-
6	UWSTS0322	9.7	52	102	51	0	0	0.01	2	-
6	SSR01012	12.6	52	102	50	0	1	0.04	2	-
6	SSR15245	13.3	49	103	49	0	4	0.12	2	-
6	SSR19672	16.2	47	98	48	0	12	0.06	2	-
6	SSR07198	17.3	51	100	49	0	5	0.04	2	-
6	SSR16020	18.6	50	105	49	0	1	0.19	2	-
6	UW026722	26.6	0	0	39	134	32	0.42	1	-
6	SSR14061	28.9	38	104	46	0	17	2.81	2	-
6	UW083805	29.9	48	110	45	0	2	1.51	2	-
6	UW025975	32.8	45	111	49	0	0	1.57	2	-
6	SSR17023	43.8	49	104	48	0	4	0.25	2	-
6	SSR13996	47.1	49	103	50	0	3	0.09	2	-
6	SSR14652	48.0	50	100	50	0	5	0.00	2	-
6	UW000036	48.6	46	93	50	0	16	0.22	2	-
6	SSR15492	49.0	46	97	53	0	9	0.52	2	-
6	DM0071	51.2	52	100	48	0	5	0.16	2	-
6	SSR18443	53.0	55	103	44	0	3	1.28	2	-
6	SSR00126	63.7	56	98	50	0	1	0.67	2	-
6	SSR14859	69.7	48	106	49	0	2	0.41	2	-
6	SSR10740	73.3	49	102	49	0	5	0.08	2	-
6	SSR19842	75.1	53	102	47	0	3	0.38	2	-
6	UWSTS0295	76.2	56	98	47	0	4	0.93	2	-
6	UWSTS0299	77.3	52	106	46	0	1	0.67	2	-
6	UWSTS0296	77.3	52	106	46	0	1	0.67	2	-

CHR	Locus	Position[z]	A/A	A/B	B/B	A/- [y]	No Data	χ^2	Df	Sig.[x]
6	UWSTS0263	77.3	49	104	46	0	6	0.50	2	-
6	UWSTS0297	77.4	50	105	46	0	4	0.56	2	-
6	SSR17604	77.6	51	101	47	0	6	0.21	2	-
6	UWSTS0304	80.0	52	103	46	0	4	0.48	2	-
6	UWSTS0302	80.0	54	103	46	0	2	0.67	2	-
6	UWSTS0303	80.0	52	105	46	0	2	0.60	2	-
6	UWSNP0125	80.0	54	105	46	0	0	0.75	2	-
6	UWSTS0266	80.0	54	105	46	0	0	0.75	2	-
6	SSR00584	80.0	54	104	46	0	1	0.71	2	-
6	UW020717	80.6	47	92	42	0	24	0.33	2	-
6	UWSTS0269	83.0	49	108	46	0	2	0.92	2	-
6	UWSTS0310	84.0	51	104	46	0	4	0.49	2	-
6	SSR18956	84.7	47	101	42	0	15	1.02	2	-
6	SSR06240	90.8	53	95	45	0	12	0.71	2	-
6	SSR16683	91.8	54	102	46	0	3	0.65	2	-
6	SSR18669	92.2	54	101	43	0	7	1.30	2	-
6	SSR17408	97.3	52	99	52	0	2	0.12	2	-
6	SSR18251	99.1	50	95	54	0	6	0.57	2	-
6	SSR03357	100.5	50	96	55	0	4	0.65	2	-
7	UWSTS0250	0.0	50	99	50	0	6	0.01	2	-
7	SSR00015	15.5	50	102	50	0	3	0.02	2	-
7	UW084483	16.6	51	102	51	0	1	0.00	2	-
7	SSR00890	17.2	52	101	47	0	5	0.27	2	-
7	SSR04689	19.1	50	106	46	0	3	0.65	2	-
7	SSR00931	20.3	53	104	48	0	0	0.29	2	-
7	SSR18648	20.6	53	105	47	0	0	0.47	2	-
7	UW083819	21.5	50	100	46	0	9	0.24	2	-
7	UW085202	21.8	51	99	47	0	8	0.17	2	-
7	UW060272	23.0	45	102	45	0	13	0.75	2	-
7	SSR07473	23.9	48	105	48	0	4	0.40	2	-
7	UW084146	24.1	45	106	48	0	6	0.94	2	-
7	UW085407	26.3	46	102	49	0	8	0.34	2	-
7	SSR11742	32.1	50	98	54	0	3	0.34	2	-
7	UW014906	33.2	50	99	55	0	1	0.42	2	-
7	SSR04704	34.6	48	101	56	0	0	0.67	2	-
7	SSR00048	42.0	49	99	57	0	0	0.86	2	-
7	SSR13885	43.3	46	100	54	0	5	0.64	2	-
7	SSR06349	44.3	46	99	55	0	5	0.83	2	-
7	UW084414	46.5	50	100	55	0	0	0.37	2	-
7	UW015467	46.5	50	100	54	0	1	0.24	2	-

[z]Position is measured in centimorgans as calculated using the Kosambi map function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001).

[y]A/- marker class includes both A/A and A/B genotypes that were indistinguishable by size-fractionation in 9% denaturing polyacrylamide gel electrophoresis.

[x]Segregation distortion at each marker locus was evaluated by the chi square test with incremental levels of significance: * = 0.10, ** = 0.05, and *** = 0.01.

Table 4. Identification and location of intervals exceeding 3 Mb in the linkage map without marker coverage. The linkage map was constructed for a *C. sativus* $F_{2:3}$ population consisting of 205 F_2 individuals derived from a cross of parthenocarpic inbred line '2A' and non-parthenocarpic inbred line 'Gy8'. All SSR, STS, and dCAPS markers screened for polymorphisms between the parental lines were analyzed with *in silico* PCR using the Gy14 draft genome assembly version 1.0 (Yang et al., 2012). The table presents the number of markers identified by *in silico* PCR as being located within the intervals unaccounted for by the linkage map. None of the markers in these intervals were found to be polymorphic in the 2A×Gy8 population. It should be noted that *in silico* PCR was only able to confirm PCR amplicons for 77.6% of markers screened by this study. Consequently, the number of markers screened within these intervals is likely greater, but the data presented provides an estimate of marker availability in these regions.

Interval[z]	Flanking Marker	Assembly Position[y]	Flanking Marker	Assembly Position[y]	Interval Length[x]	Markers Screened[w]
CHR 1-1	SSR05793	3.14	UW045607	10.25	7.11	124
CHR 1-2	UW083821	14.69	UW074644	23.69	9.01	165
CHR 1-3	UW074644	23.69	End	28.50	4.80	40
CHR 2-1	Start	0.00	UW084907	5.81	5.81	81
CHR 2-2	UW036707	18.02	UW016354	21.78	3.77	58
CHR 3-1	SSR03409	9.39	SSR07220[v]	15.82	6.43	39
CHR 3-2	SSR11397	31.86	SSR30236	35.49	3.63	42
CHR 5-1	Start	0.00	UW084492	7.03	7.03	226
CHR 5-2	UW059902	19.54	SSR13409	26.31	6.77	243
CHR 6-1	UW025975	10.00	SSR17023	14.82	4.82	52
CHR 7-1	UW015467	14.36	End	19.34	4.98	92

[z]Intervals in the linkage map must be larger than 3 Mb for inclusion in this table. Intervals are numbered in descending order beginning form the start of the Gy14 chromosome assembly version 1.0 (Yang et al., 2012).

[y]Physical position measured in Mb in the Gy14 draft genome assembly version 1.0.

[x]Measured in Mb.

[w]Number of screened markers with identifiable *in silico* PCR amplicons located within each linkage map interval.

[v]Nearest interval flanking marker with assembly position data available.

Table 5. Scaffold and assembly position data obtained from Gy14 and 9930 draft genome assemblies for the linkage map generated for a 2A×Gy8 $F_{2:3}$ population developed for the study of parthenocarpic fruit set in *C. sativus* (Huang et al., 2009; Yang et al., 2012). Genome assembly position data is only presented for 'Gy14' for simplicity, as genome assembly positions for '9930' are similar in order.

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
1	UW085383	0.0	1	171432	scaffold01141	171188	scaffold000015	172636
1	SSR13109	1.0	1	1319904	scaffold01357	853121	scaffold000015	1318239
1	SSR15108	1.7	1	1491872	scaffold01357	1025089	scaffold000015	1487555
1	SSR04644	4.3	1	2036762	scaffold01357	1569979	scaffold000015	2025931
1	SSR04304	4.9	1	2172420	scaffold01357	1705637	scaffold000015	2156216
1	SSR11654	9.0	1	2979557	scaffold00953	587667	scaffold000013	573269
1	SSR05793	10.6	1	3137574	scaffold00953	745684	scaffold000013	728957
1	UW045607	32.7	1	10251706	scaffold01365	518799	scaffold000021	1126631
1	SSR15755	37.3	1	12319513	scaffold02337	15337	scaffold000175	165308
1	UW084360	39.7	1	12962762	scaffold00485	1893	scaffold000077	711237
1	UW083897	40.3	No hit	No hit	scaffold00674	16771	scaffold000165	14872
1	SSR14526	40.7	No hit	No hit	Multiple	Multiple	scaffold000052	345728
1	UW084542	41.4	1	13796672	scaffold02805	30312	scaffold000172	61223
1	UW083821	44.6	1	14685637	scaffold02852	44889	scaffold000173	160204
1	UW074644	82.0	1	23693606	scaffold03577	79002	scaffold000019	1042925
2	UW084907	0.0	2	5808924	scaffold00894	1057126	scaffold000130	143494
2	SSR00204	0.9	2	5885745	scaffold00894	980138	scaffold000130	68804
2	SSR18937	3.7	2	7148763	scaffold01227	1301052		
2	SSR13532	4.6	2	7338112	scaffold01227	1111703	scaffold000046_1	1030689
2	UW059395	7.2	No hit	No hit	scaffold02604	181905	scaffold000018	286538
2	UW043178	7.5	No hit	No hit	scaffold01305	144	scaffold000018	516674
2	UW043203	7.6	No hit	No hit	scaffold01305	56009	scaffold000018	567779
2	UW085388	9.0	No hit	No hit	scaffold01305	354649	scaffold000018	861968

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
2	UW043299	9.1	No hit	No hit	scaffold01305	355193	scaffold000018	862510
2	UW084463	9.4	No hit	No hit	scaffold01305	434969	scaffold000018	941830
2	UWSTS0384	9.8	2	8456377	scaffold01142	366700	scaffold000420	8656
2	UW057528	10.6	2	8866813	scaffold02352	416499	scaffold000018	1383680
2	SSR04869	11.1	2	9248593	scaffold02352	798279	scaffold000018	1762396
2	UW085357	12.2	2	9671627	scaffold01337	308386	scaffold000018	2228796
2	UW085360	13.4	2	10074979	scaffold01337	711950	scaffold000200	130683
2	SSR04870	16.4	2	10977372	Multiple	Multiple	scaffold000339	8458
2	UW078361	31.6	2	13426491	scaffold03625	709774	scaffold000155	14958
2	UW078335	32.0	2	13476766	scaffold03625	659499	scaffold000155	65836
2	UW078088	33.0	2	14000885	scaffold03625	135380	scaffold000101	311862
2	UW053502	35.5	2	14554289	scaffold02219	336841		
2	UW082700	41.3	2	15661719	scaffold04100	652462	scaffold000042	796088
2	SSR16916	49.9	2	17904302	scaffold01144	413278	scaffold000011_2	389091
2	UW036707	52.1	2	18018307	scaffold01144	530999	scaffold000011_2	275086
2	UW083968	52.4	2	18018306	scaffold01144	530998	scaffold000011_2	275087
2	UW016354	72.4	2	21783698	scaffold00888	243879	Scaffold000017	234728
2	UW012751	83.6	2	22634128	scaffold00765	408079	scaffold000004_1	787077
2	SSR16028	88.9	2	23312515	scaffold00245	116531	scaffold000004_1	108690
2	SSR03606	91.5	No hit	No hit	scaffold03076	321059	scaffoldrepeat35022	109
3	SSR14159	0.0	3	137577	scaffold03080	2690444	scaffold000070	758239
3	SSR05312	5.8	3	2370399	scaffold03080	457622	scaffold000285	21171
3	SSR02451	6.5	3	2553442	scaffold03080	274579	scaffold000099	89684
3	UW085290	11.5	3	3981074	scaffold00540	1152534	scaffold000140	168141
3	SSR16408	15.4	3	5670176	scaffold00962	234940	scaffold000001	558637
3	SSR05891	21.9	3	8012567	scaffold02229	4620396	scaffold000001	2861187
3	SSR14725	23.8	3	7697836	scaffold02229	4935127	scaffold000001	2555389

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
3	SSR01573	25.3	3	8457079	scaffold02229	4175884	scaffold000001	3330107
3	UW055751	25.9	3	8537849	scaffold02229	4095114	scaffold000001	3413029
3	SSR03409	28.3	3	9390684	scaffold02229	3242279	scaffold000001	4232045
3	UW083723	47.8	No hit	No hit	scaffold00550	41626	scaffold000170	44266
3	UW084166	47.8	No hit	No hit	scaffold00550	19997	scaffold000170	60878
3	SSR07220	48.4	3	15824553	scaffold00429	2459555	scaffold000838	1871
3	SSR00525	53.0	3	17778290	scaffold00785	134290	scaffold000064	978216
3	SSR02068	53.1	3	18304175	scaffold01045	19563	scaffold000067	657731
3	SSR06210	54.8	3	19783541	scaffold00164	684306	scaffold000145	236106
3	UW083972	55.6	No hit	No hit	scaffold01183	26767	scaffold000004_2	1703931
3	SSR16056	56.9	3	21126750	scaffold02615	25556	scaffold000134	376586
3	SSR02132	57.0	No hit	No hit	scaffold00508	79944	scaffold000231	21266
3	UW084363	61.0	3	23591721	scaffold00529	88982	scaffold000196	8003
3	UW083944	62.5	3	24094877	scaffold01071	455228	scaffold000004_2	2302366
3	UW085394	76.2	3	28782977	scaffold03810	220755	scaffold000014	2337875
3	SSR13949	79.6	3	31297644	scaffold02047	2307129	scaffold000008	914693
3	SSR05328	83.1	3	31928315	scaffold03356	322746	Multiple	Multiple
3	SSR11397	83.2	3	31862604	scaffold03356	257035		
3	SSR30236	100.2	3	35493403	scaffold01110	405306	scaffold000079	561261
3	SSR23159	102.5	3	36771242	scaffold03356	191183	scaffold000097	19597
3	SSR10783	106.0	3	39075853	scaffold02995	1935413	scaffold000007	397558
3	SSR15312	106.5	3	39925248	scaffold02995	2784808	Multiple	Multiple
3	SSR20578	106.8	3	39247599	scaffold01081	2107159	scaffold000007	228765
3	UW084555	108.8	3	39610878	scaffold02995	2470438	scaffold000112	129379
4	SSR11074	0.0	4	2514320	scaffold00919	402744	scaffold000080	605175
4	UW083992	3.3	4	3098193	scaffold01374	390229	scaffold000133	180630
4	SSR05783	5.5	No hit	No hit	scaffold01374	174126	scaffold000075	24501

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
4	UW083734	6.4	4	3684579	scaffold00930	388582	scaffold000075	589100
4	UW084487	10.4	No hit	No hit	scaffold01936	75889	scaffold000224	7527
4	SSR05899	11.5	4	4447306	scaffold01233	536412	scaffold000380	9269
4	SSR01615	15.7	4	5343133	scaffold00931	150254	scaffold000012	2309236
4	UW084453	22.0	4	7089265	scaffold01221	495769	scaffold000012	634628
4	SSR12386	30.2	4	9559966	Multiple	Multiple	Multiple	Multiple
4	UW084379	33.2	4	10425135	scaffold00696	151855	scaffold000036	930557
4	UW083899	34.6	4	10801560	scaffold00696	528280	scaffold000036	543595
4	SSR05415	38.1	4	11959196	scaffold00696	1685916	scaffold000115	377232
4	SSR04482	41.5	4	12355207	scaffold01029	324058		
4	UW029413	42.5	4	12536441	scaffold01029	505292	scaffold000333	2443
4	SSR13021	45.6	No hit	No hit	scaffold01029	796712		
4	SSR04649	50.0	4	14173094	scaffold03709	66549	scaffold000040	270463
4	UW084520	50.5	No hit	No hit	scaffold02541	130056	scaffold000466	2936
4	SSR02697	50.6	No hit	No hit	scaffold02541	150365	scaffold000305	14132
4	UW083971	58.5	4	14937001	scaffold01174	492945	scaffold000052	1412471
4	SSR14393	58.7	No hit	No hit	scaffold02856	143429	no hit	
4	SSR10368	59.7	4	15661383	Multiple	Multiple	Multiple	Multiple
4	UW083893	69.2	4	17301348	scaffold00614	1038971	scaffold000055	116387
4	UW083894	71.4	4	19299941	scaffold00614	662113	scaffold000148	229596
4	SSR05515	73.8	4	19608843	scaffold00614	352806	scaffold000069	270889
4	UW084851	81.2	No hit	No hit	scaffold01000	736758	scaffold000032	372649
4	SSR16498	81.9	4	21333424	scaffold03533	153033	scaffold000032	537118
4	SSR00249	82.8	4	21422529	scaffold03533	238718	scaffold000032	626428
4	SSR18551	83.2	4	21382878	scaffold03533	204744	scaffold000032	586777
4	SSR14054	84.7	4	21715535	scaffold03533	551175	scaffold000032	919434
4	SSR18559	86.0	4	21980697	scaffold03533	828957	scaffold000032	1184596

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
4	UW084518	86.9	4	22448199	scaffold02500	150992	scaffold000032	1652098
4	UW084519	87.7	4	23048518	scaffold02500	751801	scaffold000063	491003
5	UW084492	0.0	5	7031013	scaffold02004	146179	scaffold000027	2401078
5	UW085421	1.3	5	9906775	scaffold01219	6668	scaffold000037	112202
5	UW084451	2.7	5	9967221	scaffold01219	66032	scaffold000037	172662
5	SSR11264	13.9	No hit	No hit	scaffold00444	764143	scaffold000501	7356
5	SSR16110	14.9	5	14812334	scaffold00067	35683	scaffold000026	656295
5	UW084566	15.2	5	12714540	scaffold03128	604	scaffoldrepeat021233	97
5	SSR32717	15.7	5	15381825	scaffold03611	342440	scaffold000026	86804
5	UW005172	16.8	5	15982981	scaffold00438	222634	scaffold000048	513844
5	SSR07711	17.4	5	15886255	scaffold00438	124775	scaffold000048	417118
5	SSR15321	20.4	5	17033707	scaffold02581	649423	scaffold000048	1564570
5	SSR00182	25.7	5	18153119	Multiple	Multiple	Multiple	Multiple
5	UW001903	26.4	5	18746975	scaffold00154	1396000	Multiple	Multiple
5	UW059902	32.3	5	19541183	scaffold02633	484054	scaffold000243	44653
5	SSR13409	52.6	5	26314894	scaffold02023	996514	scaffold000033	464656
5	SSR02895	52.9	5	26792520	scaffold02023	533512	scaffold000033	942282
5	UW084644	53.4	5	26996501	scaffold02023	321778	scaffold000033	1146263
5	SSR19343	53.8	5	26996607	scaffold02229	321659	scaffold000033	1146369
5	SSR20897	54.7	5	27453831	scaffold01139	83822	scaffold000033	1603593
6	SSR16163	0.0	6	155062	scaffold01044	155501	Multiple	Multiple
6	SSR02021	2.2	6	1000053	scaffold01044	999882	scaffold000016	1444214
6	UWSTS0316	3.3	6	1043302	scaffold01044	1043302	scaffold000016	1403337
6	UW084474	9.3	No hit	No hit	scaffold01398	1063	scaffold000030	764610
6	UWSTS0322	9.7	6	3523431	scaffold03611	2251056	scaffold000030	772429
6	SSR01012	12.6	6	3908627	scaffold03611	1866056	scaffold000030	1152296
6	SSR15245	13.3	6	4095588	scaffold03611	1679221	scaffold000030	1339447

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
6	SSR19672	16.2	6	5007628	scaffold03611	766859	scaffold000058	426364
6	SSR07198	17.3	6	5338298	Multiple	Multiple	Multiple	Multiple
6	SSR16020	18.6	6	5501822	scaffold03611	272461	Multiple	Multiple
6	UW026722	26.6	6	7964646	scaffold00998	2935238	scaffold000028	398565
6	SSR14061	28.9	6	8549350	scaffold00998	2350667	scaffold000024	1981847
6	UW083805	29.9	6	9019465	scaffold00998	1880419	scaffold000024	1530230
6	UW025975	32.8	6	10001697	scaffold00998	898187	scaffold000024	582672
6	SSR17023	43.8	6	14824653	scaffold00095	22439	scaffold000141	60437
6	SSR13996	47.1	6	15696372	scaffold00714	118916	scaffold000177	199018
6	SSR14652	48.0	6	15545097	scaffold02188	96209	scaffold000121	90193
6	UW000036	48.6	No hit	No hit	scaffold00013	33758	scaffold000918	3561
6	SSR15492	49.0	6	15986750	scaffold00791	117624	scaffold000121	398738
6	DM0071	51.2	6	17254742	scaffold03159	128795	scaffold000051	210278
6	SSR18443	53.0	6	17824244	scaffold01139	358725	scaffold000144	272172
6	SSR00126	63.7	6	20657022	scaffold01274	667629	scaffold000025	925793
6	SSR14859	69.7	6	21809123	scaffold01001	130173	Multiple	Multiple
6	SSR10740	73.3	6	22833658	scaffold00858	259846	Multiple	Multiple
6	SSR19842	75.1	6	23060872	scaffold00858	489483	scaffold000044	1244867
6	UWSTS0295	76.2	6	23184991	scaffold00858	623457	scaffold000044	1368986
6	UWSTS0299	77.3	6	23488310	scaffold00927	254385	scaffold000116	247822
6	UWSTS0296	77.3	6	23330171	scaffold00927	96246	scaffold000359	3983
6	UWSTS0263	77.3	6	23380051	scaffold00927	146286	scaffold000116	343810
6	UWSTS0297	77.4	6	23432714	scaffold00927	198789	scaffold000116	292681
6	SSR17604	77.6	6	23515890	scaffold00927	282536	scaffold000116	221258
6	UWSTS0304	80.0	6	23903299	scaffold00927	669374	scaffold000035	1459530
6	UWSTS0302	80.0	6	23810592	scaffold00927	576667	scaffold000035	1548837
6	UWSTS0266	80.0	6	23826817	scaffold00927	593004	scaffold000035	1532496

2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
6	UWSTS0303	80.0	6	23865714	scaffold00927	631789	scaffold000035	1494471
6	UWSNP0125	80.0	6	23988601	scaffold00927	754676	scaffold000035	1372031
6	SSR00584	80.0	No hit	No hit	Multiple	Multiple	scaffold000035	1559134
6	UW020717	80.6	6	24192119	scaffold00927	958194	scaffold000035	1165156
6	UWSTS0269	83.0	6	24733628	scaffold00927	1499827	scaffold000035	627785
6	UWSTS0310	84.0	6	24908072	scaffold00927	1674147	scaffold000035	459282
6	SSR18956	84.7	6	25028967	scaffold00927	1795223	scaffold000035	341062
6	SSR06240	90.8	6	26905762	scaffold00542	387613	scaffold000002	3224274
6	SSR16683	91.8	6	27192691	scaffold00542	674673	scaffold000002	2933762
6	SSR18669	92.2	6	27641107	scaffold00542	1121556	scaffold000002	2498111
6	SSR17408	97.3	6	28908733	scaffold00542	2390107	Multiple	Multiple
6	SSR18251	99.1	6	29283529	scaffold00542	2764834	scaffold000002	870008
6	SSR03357	100.5	6	29904596	scaffold00542	3386663	scaffold000002	261513
7	UWSTS0250	0.0	7	1238608	scaffold00926	164069	Multiple	Multiple
7	SSR00015	15.5	7	3605198	scaffold03443	84223	scaffold000010	703910
7	UW084483	16.6	7	3745275	scaffold01658	49693	scaffold000010	843987
7	SSR00890	17.2	7	3992719	scaffold01658	326543	scaffold000010	1091431
7	SSR04689	19.1	7	4563241	scaffold01658	908827	scaffold000010	1661953
7	SSR00931	20.3	7	5122138	Multiple	Multiple	scaffold000010	2220850
7	SSR18648	20.6	7	5014213	scaffold01658	1364846	scaffold000010	2112925
7	UW083819	21.5	7	6491902	scaffold02669	54787	scaffold000125	650080
7	UW085202	21.8	7	6041622	scaffold02205	136301	scaffold000125	199774
7	UW060272	23.0	7	5731331	scaffold02639	232701	scaffold000191	110543
7	SSR07473	23.9	7	6754293	scaffold03306	324467	scaffold000023	241289
7	UW084146	24.1	No hit	No hit	scaffold00284	86019	scaffold000024	2167443
7	UW085407	26.3	7	7658555	scaffold00089	298129	scaffold000023	1145510
7	SSR11742	32.1	3	29525078	scaffold02047	534563	scaffold000208	30750

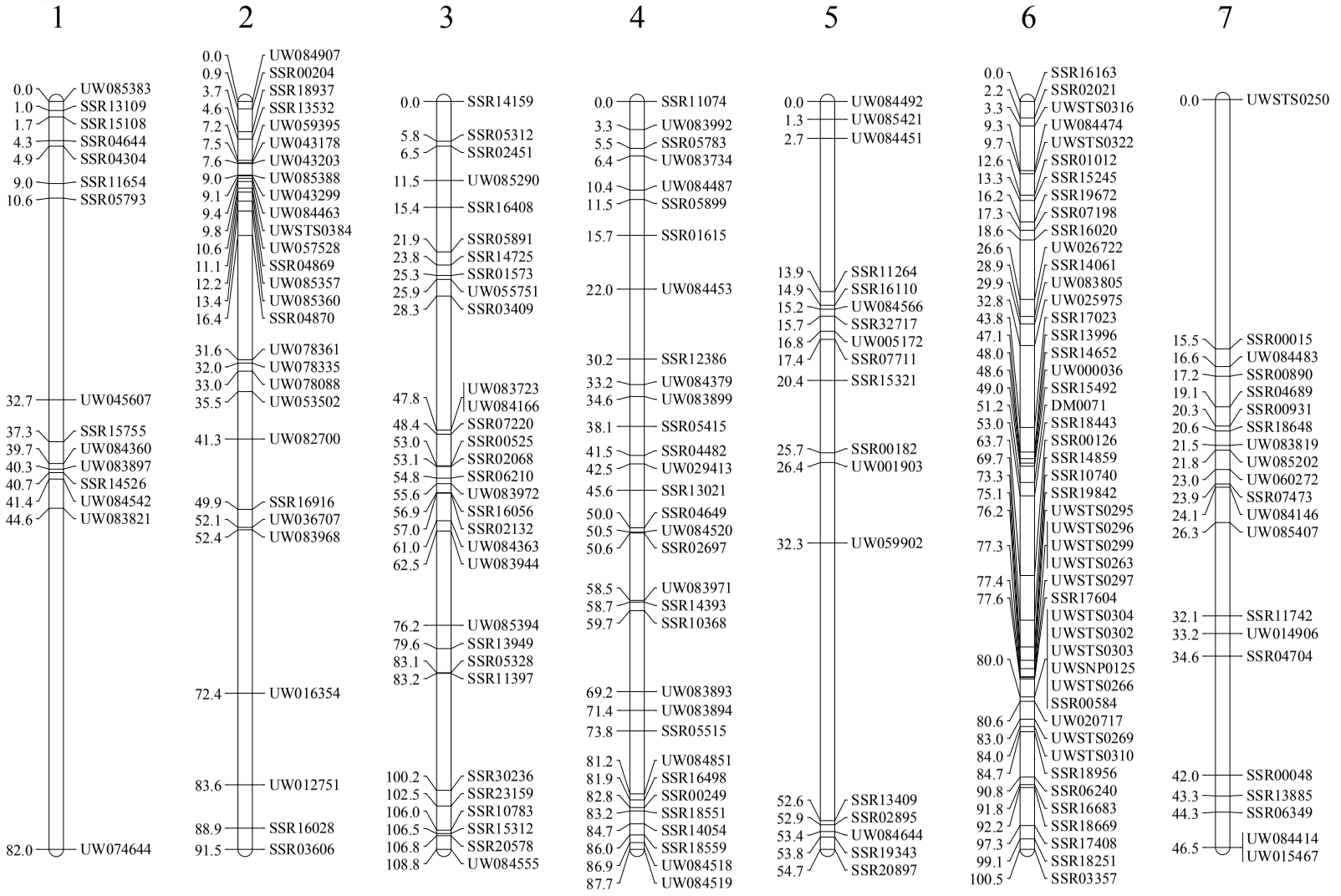
2AxGy8 Linkage Map			Gy14 Draft Genome Assembly				9930 Draft Genome Assembly	
CHR	Marker Name	Pos[z]	CHR	Fwd Pos[y]	Scaffold	Pos[x]	Scaffold	Pos[x]
7	UW014906	33.2	7	9737238	scaffold00793	2878345	scaffold000022	2548382
7	SSR04704	34.6	7	9737055	scaffold00793	2878528	scaffold000022	2548568
7	SSR00048	42.0	7	12121718	scaffold00793	493865	scaffold000022	173283
7	SSR13885	43.3	7	12684121	scaffold03967	68039	scaffold000139	275164
7	SSR06349	44.3	7	13171763	Multiple	Multiple	scaffold000199	60569
7	UW084414	46.5	7	14322278	scaffold00994	22935	scaffold000081	470573
7	UW015467	46.5	7	14355975	scaffold00862	1429	scaffold000081	504270

[z]Position measured in centimorgans and calculated using the Kosambi map function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001).

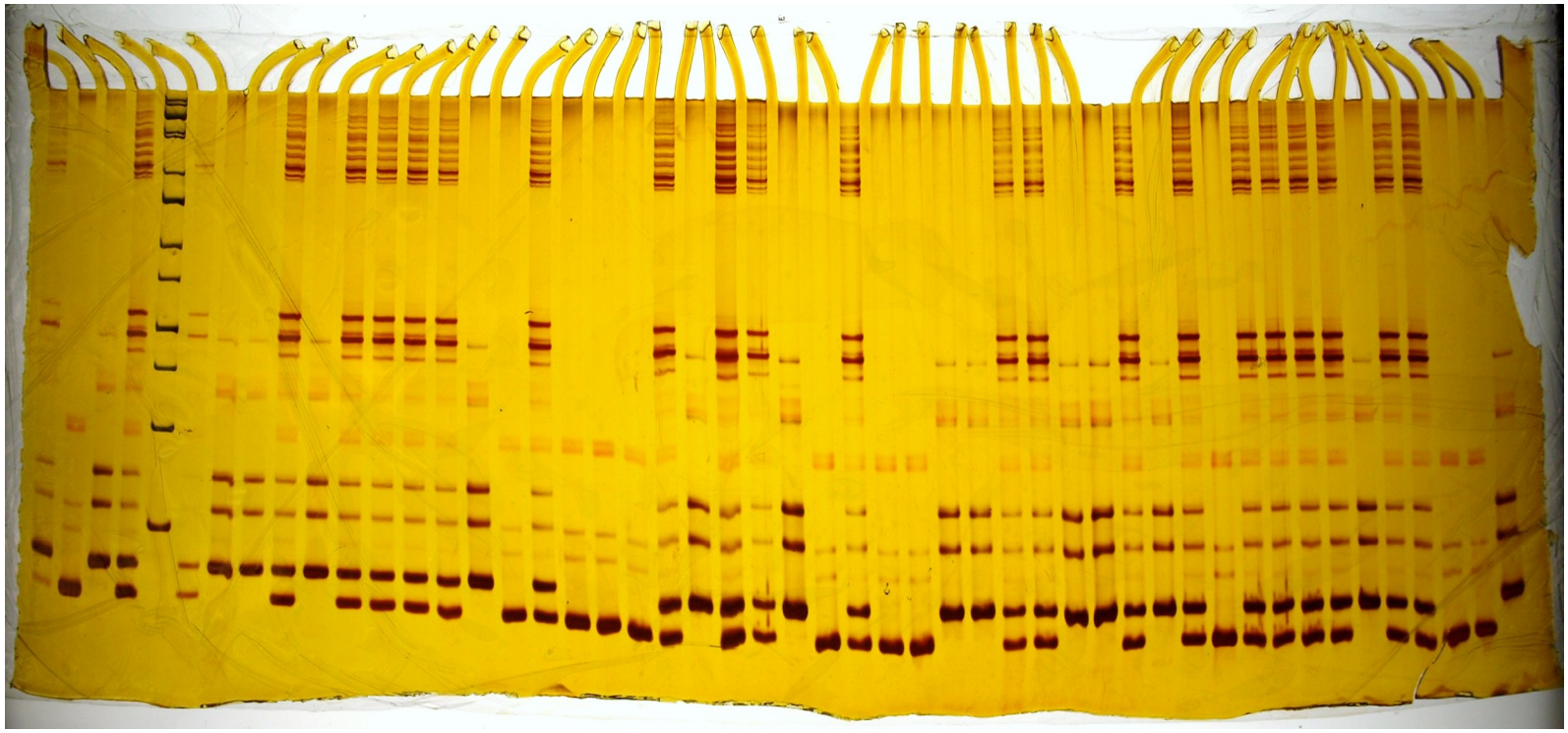
[y]Starting location of the forward primer in the Gy14 draft genome assembly measured in base pairs.

[x]Starting location of the forward primer in each scaffold measured in base pairs.

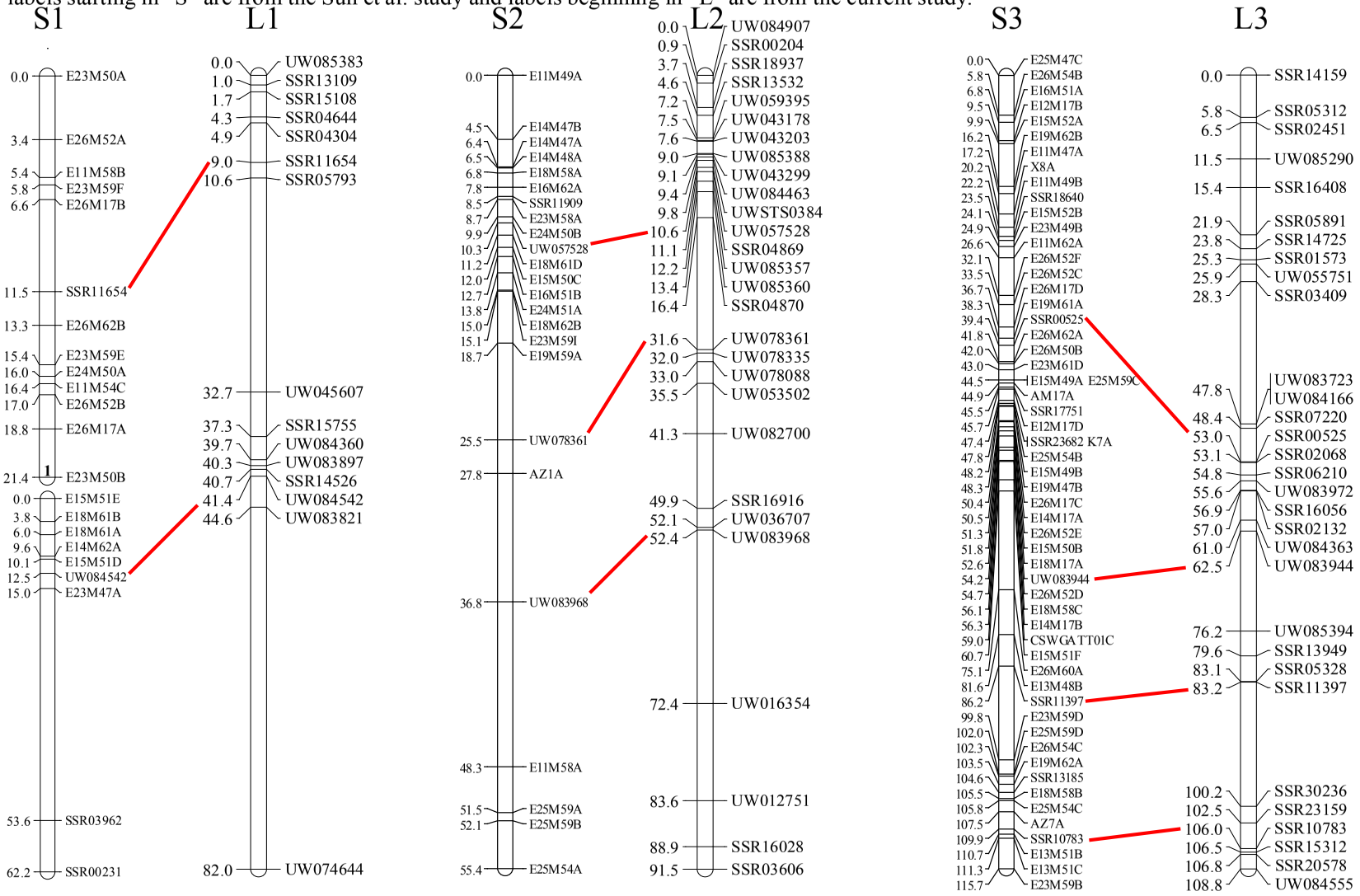
Figure 1. Linkage map constructed for a $F_{2:3}$ mapping population that was generated from the cross of a highly parthenocarpic inbred line, '2A', with a non-parthenocarpic inbred line, 'Gy8'. The resulting linkage map was constructed with 205 2A×Gy8 F_2 individuals and consists of 185 SSR, 5 STS, and 2 dCAPS marker loci in seven linkage groups covering 571.7 cM. Map distances were calculated using Kosambi's map function in JoinMap 3.0 software (Van Ooijen and Voorrips, 2001). Linkage groups were assigned to chromosomes according to Yang et al. (2012).



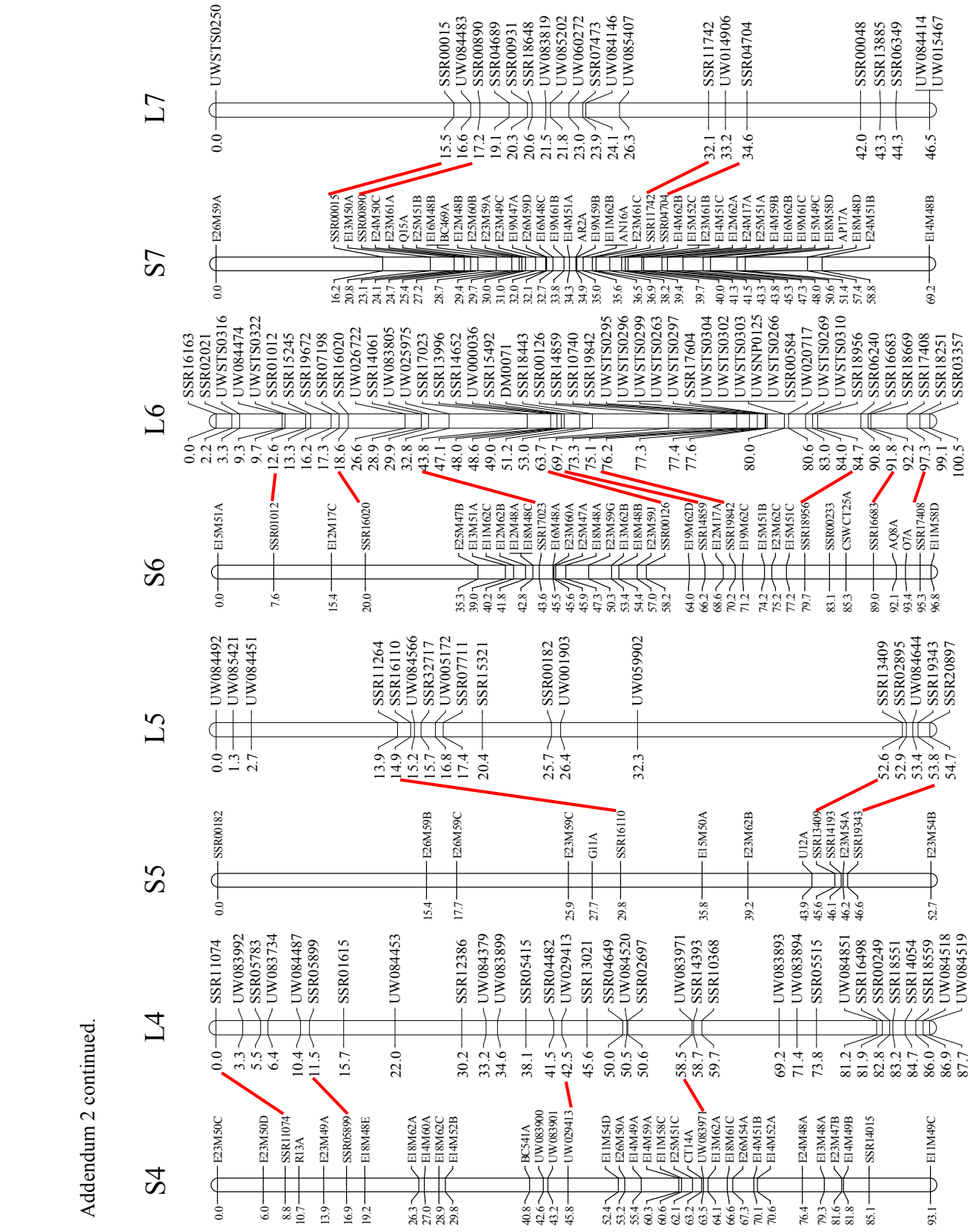
Addendum 1. Photograph demonstrating the generation of SSR molecular marker data by size-fractionation of PCR amplicons using polyacrylamide gel electrophoresis (PAGE) and silver staining. The PCR amplicons in this example were amplified with the SSR marker SSR10783. Lane 1 contains PCR amplicons amplified from parthenocarpic parent inbred line '2A'. Lane 2 contains PCR amplicons amplified from non-parthenocarpic parent inbred line 'Gy8'. Lane 5 contains GeneRuler 100 bp PLUS DNA ladder. All other lanes contain PCR amplicons amplified from 2A×Gy8 F_2 individuals in numerical order. Data in this example was collected from the major polymorphic bands located below 100bp of the DNA ladder (lowest band of the DNA ladder).



Addendum 2. Linkage map constructed with genotypic data obtained from 89 of 126 2AxGy8 F_2 individuals presented in Sun et al. (2006). To anchor the linkage map and provide points for direct comparison to the Gy14 draft genome assembly and the 2AxGy8 linkage map constructed in this study, marker data was collected from an additional 43 SSR markers. The combined linkage map consists of 236 marker loci contained in eight linkage groups. The eighth linkage group is the result of a split of the linkage group corresponding to Chromosome 1. All chromosome labels starting in "S" are from the Sun et al. study and labels beginning in "L" are from the current study.



Addendum 2 continued.



Chapter 3

Identification of Quantitative Trait Loci Associated with Parthenocarpic Fruit Set in Processing Cucumber (*Cucumis sativus* L.)

Abstract

Successful parthenocarpic cucumber cultivars have been developed and used for many years. However, the genetic inheritance of parthenocarpic expression in cucumber has not been well characterized. Therefore, an $F_{2:3}$ population was developed for a narrow cross between a highly parthenocarpic inbred line, '2A', and a non-parthenocarpic inbred line, 'Gy8', to identify QTL associated with parthenocarpic fruit set. Seven QTL associated with parthenocarpic fruit set were detected with four QTL being identified consistently in all analyses. Consensus QTL were located on chromosome 5 at 32.3 - 54.7 cM (*parth5.1*), chromosome 6 at 0.0 - 9.7 cM (*parth6.1*), chromosome 6 at 80.0 - 83.0 cM (*parth6.2*), and chromosome 7 at 21.8 - 32.1 cM (*parth7.1*). All QTL were additive and significant epistatic interactions were not detected. The locations of these four QTL were compared with QTL identified for parthenocarpy and yield in previous studies. Yield QTL were not found to co-localize with the QTL identified for parthenocarpic fruit set in this study. In addition to parthenocarpic expression, seed size and seed weight traits were observed to segregate in this population. Seed size and seed weight traits were highly correlated and QTL analyses of both traits revealed similar results. Two QTL consistently associated with seed size and weight were identified on chromosome 5 at 14.9 - 20.4 cM and on chromosome 6 at 80.0 cM. The QTL on chromosome 6 at 80.0 - 83.0 cM was identified as being a potentially pleiotropic locus affecting both parthenocarpic expression and seed size and weight.

Sequence information was extracted from the genomic region encompassed by each of the four QTL associated with parthenocarpic fruit set and explored for Arabidopsis gene homologs with the BLASTn tool provided by NCBI. Multiple candidate genes were identified and potentially promising candidates are discussed in depth.

Introduction

Parthenocarpy is a desirable trait for the production of fruit and vegetable crops that have undesirable large and/or hard seeds. Parthenocarpic cultivars are also beneficial in the production of crops in which pollination is difficult or impacted by abiotic factors and these cultivars often result in increased yield. Naturally occurring parthenocarpy has been observed in many plant species and parthenocarpic cultivars are common in citrus, cucurbit, and solanaceous crop production (Beraldi et al., 2004; Fos et al., 2000; Gorguet et al., 2008; de Menezes et al., 2005; Miyatake et al., 2012; Sun et al., 2006; Vardi et al., 2008). Models of inheritance vary widely from simple inheritance to complex quantitative inheritance within each of these species depending upon population materials (Beraldi et al., 2004; Fos et al., 2000; Gorguet et al., 2008; de Menezes et al., 2005; Miyatake et al., 2012; Sun et al., 2006; Vardi et al., 2008). There are disagreements among studies on the mode of genetic inheritance in cucumber. Hawthorn and Wellington (1930) and Meshcherov and Juldasheva (1974) both reported models consisting of a single recessive gene for the inheritance of parthenocarpy. Pike and Peterson (1969) also developed a single gene model and reported parthenocarpy to be inherited as a single incompletely dominant gene. Kvasnikov et al. (1970) were the first to propose complex inheritance for parthenocarpy with a model consisting of many recessive genes. This was

followed by a proposal by de Ponti and Garretson (1976) of an additive three gene inheritance model. Similarly, El-Shawaf and Baker (1981) found parthenocarpy to be quantitatively inherited with both additive and non-additive gene effects. Most recently, Sun et al. (2006) reported four major genomic regions associated with parthenocarpic expression with significant epistasis and large genotype×environment interactions.

Genes associated with natural parthenocarpy have not been cloned and the events directly involved in the initiation of fruit set remain unknown. However, reports from Arabidopsis and tomato consistently report elevated expression of gibberellins in association with parthenocarpic expression (Dorcey et al., 2009; Fos et al., 2000; Olimpieri et al., 2007; Pascual et al., 2009; Serrani et al., 2007; Serrani et al, 2008; Serrani et al., 2010). Interestingly, both auxin and cytokinin induced parthenocarpic expression have been found to be mediated by gibberellins and gibberellin biosynthetic inhibitors can block auxin and cytokinin induced parthenocarpic expression (Ding et al., 2013; Fuentes et al., 2012; Serrani et al., 2008; Serrani et al., 2010). These observations demonstrate a key role for gibberellins in parthenocarpic expression and fruit set.

Parthenocarpy has also been exploited in crop production through the application of exogenous phytohormones (Gillaspy et al., 1993; Vivian-Smith and Koltunow, 1999). Auxin, gibberellic acid, cytokinin, and combinations of these are the most common phytohormones used to induce parthenocarpic expression (Pandolfini, 2009). In addition to auxin, gibberellic acid, and cytokinin, the exogenous application of brassinosteroids was found to induce parthenocarpic fruit set in cucumber (Fu et al., 2008). Although not previously associated with parthenocarpic expression, the exogenous application of brassinosteroids has been used to increase yields in crop production (Divi and Krishna, 2009; Vriet et al., 2012). The ability to induce

parthenocarpic expression with multiple hormones exemplifies the complexity behind hormone signaling and illuminates the paradox of how eight known plant hormones (auxins, gibberellins, cytokinins, ethylene, abscisic acid, brassinosteroids, jasmonic acids, and strigolactones) can regulate every physiological process in plants.

Brassinosteroids are a large family of growth promoting polyhydroxylated steroid hormones involved in regulating numerous aspects of physiological response during vegetative and reproductive development (Clouse, 2011). Brassinolide and its precursor, castasterone, are the most commonly found brassinosteroids in plants (Srivastava, 2002). After the discovery of brassinosteroids during the 1970's, numerous molecular genetic and biochemical studies were conducted utilizing brassinosteroid deficient and insensitive mutants to elucidate the brassinosteroid signaling pathway. Brassinosteroids are perceived by the brassinosteroid receptor Brassinosteroid Insensitive 1 (BRI1). BRI1, a leucine-rich-repeat containing receptor-like kinase (LRR-RLK), acts as a transmembrane brassinosteroid receptor for the brassinosteroid signaling pathway (Li and Chory, 1997). Binding of BRI1 with brassinosteroids activates the intracellular domain of BRI1 through phosphorylation and allows for association with its co-receptor, BRI1-Associated Receptor Kinase 1 (BAK1), which further enhances the kinase activity of BRI1 (Li et al., 2002; Nam and Li, 2002; Wang et al., 2008). Activated BRI1 leads to a number of intermediate phosphorylation and dephosphorylation steps before ending with two transcription factors, Brassinazole Resistant 1 (BZR1) and Brassinazole Resistant 2 (BES1), which regulate brassinosteroid responsive gene expression (He et al., 2005; Sun et al.; 2010; Yin et al, 2005; Yu et al., 2011).

The objective of this research is to determine a model of inheritance and identify quantitative trait loci (QTL) associated with parthenocarpic fruit set in cucumber. In order to

accomplish this objective, a new approach to phenotypically evaluating parthenocarpic potential focusing on early fruit initiation and development in cucumber was utilized. Traditional QTL mapping approaches such as interval mapping (IM), composite interval mapping (CIM), and multiple interval mapping (MIM) were employed to detect and construct optimal models for the inheritance of parthenocarpic fruit set. After identification of a consensus QTL model, genomic regions associated with parthenocarpic fruit set were explored for potential candidate genes. Finally, a candidate gene model for future investigation is proposed and QTL for use in increasing efficiency by breeding programs seeking to incorporate parthenocarpic expression into elite cucumber breeding populations are presented.

Materials and Methods

Mapping Population

An $F_{2:3}$ mapping population was created for identification of QTL associated with parthenocarpic fruit set from a cross between the highly parthenocarpic processing cucumber inbred line, '2A', and the non-parthenocarpic processing cucumber inbred line, 'Gy8' (Chapter 1 Figure 6). The population consists of 205 F_3 families. Phenotypic data was collected from 11 plants from each F_3 family with the exception of four F_3 families which were represented by only 6 F_3 individuals. Each F_3 plant was scored for the number of ovaries initiating parthenocarpic fruit set (Chapter 1). For construction of a genetic linkage map, the mean value obtained for each F_3 family was assigned as the phenotype of the F_2 plant from which it was derived. A genetic linkage map consisting of 185 SSR, 5 STS, and 2 dCAPS marker loci was constructed with genotypic data collected from all 205 F_2 individuals (Chapter 2).

In addition to parthenocarpic fruit set, a difference in seed size and weight between the parental lines was observed as a segregating secondary trait in this population (Addendum 1). Parental inbred line '2A' was observed to have smaller seeds than parental inbred line 'Gy8'. Seed size was measured as seed length multiplied by seed width. Seed length was measured as the length (cm) of five healthy and fully developed seeds laid end to end. Seed width was measured as the width (cm) of five healthy and fully developed seeds laid side by side. Seed weight was measured as the total weight in grams of 50 healthy and fully developed seeds. The seeds measured in this population were seeds obtained by self-pollination of each of the 205 2A×Gy8 F_2 plants of the experimental population used for the study of parthenocarpic fruit set.

QTL Mapping of Parthenocarpic Fruit Set

All QTL analyses were performed with the statistical software R version 3.0.2 with the QTL mapping package "qtl" (R/qtl) version 1.30.4 (Broman et al., 2003). QTL analyses were performed with datasets consisting of data collected from experiment 1 alone, experiments 2 and 3 combined, and the pooling of experiments 1-3. A detailed discussion for the justification of these datasets is in Chapter 1. Briefly, the decision to analyze experiment 1 individually was due to observed differences in the timing of fruit set and the change in the experimental design that was implemented in experiments 2 and 3 after crowding of plants was observed in experiment 1. The construction of a dataset consisting of experiments 2 and 3 combined was made because of the high level of similarity between the data from each experiment. Finally, a Spearman rank correlation performed between the data of experiments 1-3 determined that although differences were observed in experiment 1, the rank of F_3 family means between all three experiments were

positively correlated and thus a dataset could be constructed with data from experiments 1-3 pooled.

Differences in the timing and location of fruit set were observed between the parental lines and this was further explored. In order to do this, the data from experiments 2 and 3 were modified to only include parthenocarpic fruits initiating development on the first 10 and first 20 nodes (5 and 15 scorable nodes, respectively, due to the trimming of the bottom 5 nodes) of each F_3 plant. Data from experiment 1 was not included in these analyses due to the noted delay in fruit set related to crowding. QTL analyses were performed on the datasets consisting of data from the first 10 and first 20 nodes of experiments 2 and 3.

Each dataset was analyzed by IM, CIM, and MIM QTL detection approaches. With the MIM approach, QTL and potential interactions between QTL were preliminarily evaluated with the *scantwo* function. The analysis was conducted with Haley-Knott regression. A permutation test with 1000 replications was used to determine LOD significance thresholds at alpha = 0.05 and 0.10 levels. With the output of the *scantwo* function providing a general view of the major QTL that would be included in a best fit QTL model for parthenocarpic fruit set, the function *stepwiseqtl* was used to perform a forward and backward search to identify the best fit QTL model. The *stepwiseqtl* function utilizes penalized LOD scores to evaluate the addition of each QTL or interaction term added to the QTL model (Manichaikul et al., 2009). The penalized LOD score seeks to control the false positive discovery rate by using a penalty to keep the rate of inclusion for spurious QTL at a predefined level (Manichaikul et al., 2009). The LOD thresholds calculated via *scantwo* are used to calculate the penalties via the *calc.penalties* function. In model selection, the *stepwiseqtl* algorithm first performs a single QTL genome scan to identify the QTL position with the largest LOD score. Then a scan for additional additive QTL,

interacting QTL, and possible pairwise interactions between QTL is performed. After the addition of each QTL, the position of each QTL is refined while taking into account the positions of the other QTL in the model. The addition of QTL terms to the model continues to a predefined maximum threshold. The backward elimination step then considers removing the smallest QTL or QTL interaction term one at a time all the way back to the first term of the model. Finally, the model with the highest overall penalized LOD score and least number of terms is selected. Once the QTL terms of the best fit model are selected, the functions *makeqtl* and *fitqtl* are used to construct the model and evaluate the fit of the model. The function *lodint* was used to determine 1.5 LOD confidence intervals for each QTL.

IM was conducted utilizing Haley-Knott regression and the *scanone* function. A permutation test with 1000 replications was used for each dataset to determine LOD significance thresholds at alpha = 0.05 and 0.10 levels. All significant QTL identified were assembled into a best fit QTL model using the functions *makeqtl* and *fitqtl*. The function *lodint* was used to determine 1.5 LOD confidence intervals for each QTL.

CIM was conducted utilizing Haley-Knott regression and the *cim* function. The number of marker covariates selected is critical to the accuracy of CIM and the number of marker covariates should ideally reflect the number of true QTL. In the preliminary evaluation of each dataset with the CIM approach, the number of marker covariates was set to be equal to the number of QTL detected by the IM approach. A permutation test with 1000 replications and a specified number of covariates was used for each dataset to determine LOD significance thresholds at alpha = 0.05 and 0.10 levels. For each dataset, an initial QTL analysis was performed with the window size set at 10 cM and the number of marker covariates set to the number of QTL observed in IM. A plot of LOD curves was then produced and inspected for the

detection of additional QTL. Evidence for inclusion of additional QTL was concluded when the CIM LOD curve extended above the alpha = 0.05 threshold a greater number of times than the number specified marker covariates. In this case the number of marker covariates was redefined to reflect the new number of expected QTL. A new permutation test was performed with the new appropriate number of marker of covariates specified. A new QTL analysis was performed and the process repeated until the number of expected QTL was in agreement with the number of specified marker covariates. All significant QTL identified were assembled into a best fit QTL model using the functions *makeqtl* and *fitqtl*. The function *lodint* was used to determine 1.5 LOD confidence intervals for each QTL.

QTL Mapping of Seed Size and Seed Weight

QTL analyses of seed size and seed weight were performed with datasets consisting of data collected from the seeds obtained by self-pollination of 205 2A×Gy8 F_2 plants. Each dataset was analyzed by MIM, CIM, and IM QTL detection approaches. With each approach, the QTL analysis was conducted identically for seed size and seed weight as it was described for parthenocarpic fruit set. All QTL analyses were performed with R/qtl.

Identification of Candidate Genes

After identification of four QTL associated with parthenocarpic fruit set that had consensus among the five datasets, the genome sequence surrounding each QTL was explored for candidate genes. For each QTL, the genome sequence between the flanking markers of a 1.5 LOD confidence interval was extracted from the Gy14 Draft Genome Assembly Version 1.0 (Yang et al., 2012). This sequence was imported into the nucleotide-nucleotide Basic Local Alignment

Search Tool (BLASTn) version 2.2.29+ utility customized for plant genomes provided by NCBI (Altschul et al., 1997). All BLASTn default settings were used with the exception of limiting the search database to only *Arabidopsis thaliana* mRNA sequences. All sequence matches were considered and matches were of high interest if they were found to associate with matches identified in one of the other three QTL regions. Genome positions of candidate genes were identified by using the Arabidopsis CDS of each gene as a query in a BLASTn search of Gy14 draft genome assembly. After a rough match for the position of each candidate gene was identified, the entire match and an additional 5 kb of upstream and downstream flanking sequence was extracted. This extended region was structurally annotated with the FGENESH utility provided by Softberry in order to predict the structure of each candidate gene in cucumber (Solovyev et al., 2006).

Alignment of Re-sequencing Data for BRI1 and BAK1 Genes

The genes predicted by Softberry with shared identities with BRI1 and BAK1 of *Arabidopsis thaliana* found in the major QTL regions of chromosome 6 at 0.0 - 9.7 cM (*parth6.1*) and at 80.0 - 83.0 cM (*parth6.2*), respectively, were compared by sequence alignment between the parental lines. Assembled whole genome re-sequencing data obtained for the parental lines '2A' and 'Gy8' (Chapter 2) along with data from the Gy14 draft genome assembly, which was used as a reference, were used in alignment. Alignments of DNA, predicted mRNA, and predicted protein sequences for both parental lines and 'Gy14' were performed using ClustalW2 software (Larkin et al., 2007). Sanger sequencing of the parental lines was used to validate any polymorphisms identified between the parental lines in the re-sequencing data. In addition, at least five dCAPS

markers for each candidate gene were designed for SNPs identified from the re-sequencing data as validation.

Sanger Sequencing of BAK1 and BRI1

The predicted genes with shared identities with BAK1 and BRI1 of *Arabidopsis thaliana* were sequenced via Sanger sequencing of the parental lines '2A' and 'Gy8'. The predicted gene sequence for each gene was obtained from the results of the Softberry FGENESH analysis. The sequence used for sequencing began approximately 1 kb before the structural gene and ended approximately 150 bp after the polyadenylation site. The sequence was divided into overlapping sections approximately 1.5 kb in length. Primers were designed in opposite orientations at the ends of each section and served as the start sites for individual Sanger sequencing reactions. Sequencing reads were expected to extend a minimum of 800 bp and overlap at the center of each 1.5 kb segment. Primers were designed with the primer design software, Primer 3 (Rozen and Skaletsky, 2000). PCR protocols were the same as described in Chapter 2. PCR amplicons were size-fractionated in 3% agarose gel and visualized with ethidium bromide staining. The PCR amplicon band for each primer pair was cut from the agarose gel and purified with the use of a Qiaex II Gel Extraction Kit (Qiagen Sciences, Germantown, Maryland). BigDye Terminator (Applied Biosystems, Foster City, CA) sequencing reactions were used to label the DNA for Sanger sequencing. Each reaction consisted of: 1 μL of diluted DNA (25ng/μL), 1 μL of 5 μM primer, 4 μL of BigDye Terminator reaction mix, and 4 μL of water for a final reaction volume of 10 μL. The BigDye Terminator PCR program is as follows: 5 min initial denaturation at 96°C; 25 cycles of 10 s at 96°C for denaturation, 5 s at 50°C for annealing, and 4 min at 60°C for extension. Excess BigDye Terminator was removed and PCR amplicons were purified with

CleanSeq magnetic beads (Agencourt Bioscience Corporation, Beverly, MA) prior to submission to the University of Wisconsin Biotechnology Center for Sanger sequencing. Alignment of candidate gene sequence obtained via Sanger sequencing from both parental lines was performed with ClustalW2 software.

Results and Discussion

Identification of QTL Associated With Parthenocarpic Fruit Set

In order to determine the optimal QTL model for the inheritance of parthenocarpic fruit set in this 2A×Gy8 cucumber population, MIM, CIM, and IM QTL mapping approaches were utilized. The decision to utilize multiple QTL detection approaches was made in order to build confidence for the inclusion or exclusion of QTL in the optimal model. Each approach has benefits in QTL detection. The IM approach is strongest in detecting single QTL traits. Since parthenocarpic fruit set is inherited as a complex trait in the 2A×Gy8 population (Chapter 1), the IM approach is utilized here as a starting point in model construction. IM is also useful in generating LOD curves for visual inspection of data and QTL quality. LOD curves can also be generated with CIM, but the use of covariates can lead to inflated LOD scores (Broman and Sen, 2009). Similarly, LOD curves generated by MIM in R/qtl described here will reflect slightly different LOD values. Where IM and CIM calculate LOD scores by comparison of models consisting of a single QTL of interest with the null model, MIM compares full QTL models with a model consisting of the full model with the QTL of interest and all of its interaction terms omitted to calculate LOD scores (Broman and Sen, 2009). In addition, when significant epistatic

interactions between QTL are present, the model will call for inclusion of QTL that may not appear significant when considered alone.

The CIM and MIM approaches are better suited to detecting multiple QTL. CIM uses marker covariates as proxies for detected QTL. The inclusion of the marker covariates in the model removes most of the effects of the QTL, which would otherwise appear as residual variation, and thus increases the power to detect additional QTL with smaller effects (Broman and Sen, 2009). By using simultaneous consideration of multiple QTL, the MIM approach is valuable in reducing residual variation from QTL with large effects, separating linked QTL, and detecting epistatic interactions between QTL (Broman and Sen, 2009). Specifically, MIM is the only approach used by this study with the ability to search for and detect epistatic interactions between QTL.

QTL analyses of data obtained from the pooling of experiments 1-3 indicated the presence of seven unique additive QTL accounting for 73.0% (CIM) - 75.5% (MIM) of the observed phenotypic variation for parthenocarpic fruit set (Table 1). The two QTL detected on chromosome 6 at 2.2 - 9.7 cM and 80.6 - 83.0 cM together accounted for approximately 26 - 40% of the observed phenotypic variation depending upon the method used. The remaining QTL each accounted for less than 10% of the observed phenotypic variation. Epistatic interactions between QTL were not detected in this dataset. Both the CIM and MIM approaches detected the same QTL with only slight differences in location, which was due to the use of the *refineqtl* function with MIM that led to the shifting of the QTL to better fitting locations (Table 1). The appearance of a large change in position for the QTL identified on chromosome 5 is an artifact of the low marker density on the genetic linkage map for chromosome 5 as neighboring markers are located more 20 cM apart (Table 1, Chapter 2 Table 5). IM was only able to detect

five of the seven QTL, but was in agreement with the other methods on the position of those QTL (Figure 1, Figure 2, Table 1). IM failed to confirm the presence of a QTL on chromosome 2 at 0.0 cM and a third linked QTL on chromosome 6 at 53.0 cM. This is not unexpected with the IM approach as it is weaker in detection of small and/or linked QTL. All analyses indicated the presence of complementation between the parental lines of this population for the inheritance of parthenocarpic fruit set. Favorable alleles for parthenocarpic fruit set at the QTL detected on chromosome 5 at 32.3 - 54.7 cM, chromosome 6 at 2.2 - 9.7 cM, and chromosome 7 at 21.5 - 21.8 cM were obtained from the parthenocarpic parental inbred line '2A'. Favorable alleles for parthenocarpic fruit set at the QTL detected on chromosome 2 at 0.0 cM, chromosome 4 at 83.2 - 87.7 cM, chromosome 6 at 53.0 cM, and chromosome 6 at 80.6 - 83.0 cM were obtained from the non-parthenocarpic parental inbred line 'Gy8'. None of the analyses detected QTL with significant dominance effects at alpha = 0.05.

QTL analyses of data from experiment 1 indicated the presence of as many as six (MIM) unique additive QTL accounting for 62.6% (CIM) - 69.0% (MIM) of the observed phenotypic variation for parthenocarpic fruit set (Table 2). Again, CIM and MIM approaches returned similar results with the key difference being the ability of MIM to detect a significant third linked QTL located between the two major QTL on chromosome 6 (Table 2). In addition, MIM detected a significant epistatic interaction between two of the QTL located on chromosome 6 at 13.3 cM and 53.0 cM. However, this interaction only accounts for 2.9% of the observed phenotypic variation (Table 2). With this dataset, IM failed to detect the QTL on chromosomes 4 and 5 (Figure 1, Figure 2). Favorable alleles at each QTL were the same as found in the pooled data from experiments 1-3. Most importantly, when comparing the QTL analyses of the data from experiment 1 with the analyses done for experiments 1-3, all methods omit the presence of

QTL on chromosome 7 and diminish the effect of the QTL on chromosome 5 in experiment 1 (Figure 1, Figure 2, Table 1, Table 2). The effect of the QTL on chromosome 2 is approximately twice as large in experiment 1 as it in the pooled analysis, indicating that the presence of the QTL is strongly associated with the data obtained in experiment 1 (Table 1, Table 2). None of the analyses detected QTL with significant dominance effects at alpha = 0.05.

QTL analyses of the combined data obtained from experiments 2 and 3 identified the presence of four unique additive QTL accounting for 54% of the observed phenotypic variation for parthenocarpic fruit set by each of the QTL detection approaches (Table 3). MIM did not detect epistatic interactions between QTL, nor the presence of a third linked QTL between the two major QTL of chromosome 6 with this dataset. In addition, this dataset did not detect the QTL on chromosomes 2 and 4 that were present in the dataset from experiment 1 (Figure 1, Figure 2, Table 2, Table 3). CIM and MIM each placed the four QTL in identical positions with IM shifting the positions slightly (Table 3). Interestingly, all four QTL detected appear to have similar effect and contributions to observed phenotypic variation (approximately 10-15% each) (Table 3). Favorable alleles at each QTL were the same as found in the pooled data from experiments 1-3. None of the analyses detected QTL with significant dominance effects at alpha = 0.05. A comparison of the QTL analyses of the data collected from experiments 2 and 3 with the data collected in experiment 1 showed that there was disagreement on the inclusion of QTL on chromosomes 2, 4, and 7. However, the LOD score curves for the QTL detected on chromosome 4 show elevation in both analyses although it never crosses the alpha = 0.10 LOD threshold in the analyses of the combined dataset from experiments 2 and 3 (Figure 1, Figure 2). This supplies weak evidence for confirmation the QTL on chromosome 4. The QTL on

chromosome 2 is undetectable in the combined data from experiments 2 and 3. The QTL on chromosome 7 is undetectable in the analyses of data from experiment 1.

QTL analyses of datasets constructed from data collected in experiment 1 alone, experiments 2 and 3 combined, and the pooled data from experiments 1-3 all detected models consisting of four to seven QTL associated with parthenocarpic fruit set (Figure 1, Figure 2, Table 1, Table 2, Table 3). All analyses were highly consistent in the placement of detectable QTL across datasets and QTL detection methodologies. All analyses confirm the presence of QTL of moderate to large effect on chromosome 5 at 32.3 - 54.7 cM (wide range due to low marker density in this genomic region), chromosome 6 at 0.0 - 9.7 cM, and chromosome 6 at 80.0 - 83.0 cM. Due to the noted experimental issues related to plant crowding observed in experiment 1, more confidence should be placed in the QTL modeling from the combined data of experiments 2 and 3. With the QTL only detected in the analysis of experiment 1 (chromosomes 2 and 4), it is plausible that these QTL are related to parthenocarpic fruit set and/or yield in high stress environments. Similarly, the high stress environment may potentially explain the absence of the QTL from chromosome 7 in the analysis of data from experiment 1.

The presence of a third linked QTL on chromosome 6 at 53.0 cM was not detectable in the analysis of combined data from experiments 2 and 3. Inspection of the LOD curves obtained through interval mapping with datasets from experiment 1 alone and experiments 2 and 3 combined show large broad QTL peaks for the QTL on chromosome 6 centered at 80.0 - 83.0 cM (Figure 1, Figure 2). In addition, there is a slight uptick in LOD scores around 53.0 cM in both datasets, although the change in LOD score is less than 1.0 in data collected from experiments 2 and 3 (Figure 1, Figure 2). These observations indicate that an additional QTL linked to the QTL on chromosome 6 at 80.0 - 83.0 cM may be present. The analyses of the

dataset from experiment 1, where the presence of the linked QTL is detected, indicates that the possible linked QTL are in coupling phase and this may explain the large LOD scores attributed to the QTL at 80.0 - 83.0 cM (Table 2). The epistatic interaction detected between the QTL on chromosome 6 at 13.3 cM and 53.0 cM may also partially explain the detection of the linked QTL in experiment 1 (Table 2). Since this interaction was not significant in the combined dataset from experiments 2 and 3, it may have led to the failure to identify the linked QTL if this locus acts epistatically. Alternatively, the uptick in LOD scores and broad QTL peak may be related to large linkage blocks and crossover events in a few individuals in this genomic region. However, no evidence for this occurrence was observed in analysis of the genotypic data in Chapter 2.

Ultimately, the pooling of data from experiments 1-3 provides the best fitting QTL model and accounts for a very large amount of the phenotypic variation observed for parthenocarpic fruit set (73.0 - 75.5%). The data from experiment 1 remains highly valuable as QTL with moderate to large effects were detected and validated, despite the observed complications related to plant crowding. However, the QTL on chromosomes 2 and 4, which were only detectable in data from experiment 1, should be considered cautiously. At best, these two QTL can only be considered minor QTL as they each only account for approximately 5% of the observed phenotypic variation in the pooled dataset. The presence of linked QTL on chromosome 6 at 53.0 cM and 80.0 - 83.0 cM remains inconclusive and will require further marker saturation in these genomic regions and possibly validation with another population with more individuals.

In order to determine if the observed differences between the parental lines in the timing and location of fruit set would reveal unique QTL related to early parthenocarpic fruit set, QTL analyses were performed with datasets consisting of data from the first 10 and first 20 nodes of

each F_3 plant. QTL analyses of data from the first 10 and first 20 nodes of each F_3 plant were only conducted with data from experiments 2 and 3 due to the observed delay in fruit set attributed to plant crowding in experiment 1. Analysis of data from the first 20 nodes (15 scorable nodes due to the trimming of the bottom 5 nodes) indicated the presence of as many as five (MIM) additive QTL accounting for 61% (CIM) - 65% (MIM) of the observed phenotypic variation for parthenocarpic fruit set (Figure 3, Figure 4, Table 4). All analyses returned consistent results with QTL identified on chromosome 5 at 52.9 cM, chromosome 6 at 0.0 cM, chromosome 6 at 80.0 cM, and chromosome 7 at 24.1 cM (Table 4). MIM identified an additional QTL with small effect on chromosome 4 at 86.9 cM. These QTL were also consistent with those identified in the analyses of the datasets collected for the first 30 nodes of experiment 1 alone and experiments 2 and 3 combined. Again, all four QTL, excluding the QTL on chromosome 4, appear to have similar effect and contributions to observed phenotypic variation (approximately 10-15% each), except for the QTL on chromosome 6 at 80.0 cM which has twice the effect of the other QTL (Figure 3, Figure 4, Table 4). The reason for the increase in the effect of the QTL on chromosome 6 at 80.0 cM in the dataset collected for the first 20 nodes versus the dataset collected for all of the data collected for experiments 2 and 3 combined (30 nodes) is unknown. It may be a reflection of the importance of the locus to parthenocarpic fruit set in the first 20 nodes of plant growth. Alternatively, it may be related to the possibility of a second linked QTL in this region as discussed previously. However, if two linked QTL do exist in this region they were again inseparable in this dataset by all QTL detection methods examined. Favorable alleles at each QTL were the same as found in experiments 1-3 for data collected from the first 30 nodes. Interestingly, only the QTL on chromosome 4 at 86.9 cM and chromosome 6 at 80.0 cM show favorable alleles being contributed from ‘Gy8’. This result

better aligns with the expectation that favorable alleles for parthenocarpic fruit set would be contributed by the parthenocarpic parental line '2A'. MIM did not detect epistatic interactions between QTL. None of the analyses detected QTL with significant dominance effects at alpha = 0.05.

From a practical perspective, the QTL results and modeling from data collected for the first 20 nodes of plant growth were nearly identical to that collected for the first 30 nodes of plant growth in experiments 2 and 3 (Table 3, Table 4). This demonstrates that future studies will be capable of phenotypically evaluating parthenocarpic fruit set with as few as 20 nodes of plant growth. This observation also complements the fact that parthenocarpic processing cucumber lines are typically commercially harvested at approximately 20 nodes of plant growth (Chapter 1 Addendum 7). In addition, it satisfies any concern related to the ability of an individual plant to set a second flush of fruit as a confounding factor in this study (Chapter 1).

Analysis of data from the first 10 nodes (5 scorable nodes due to the trimming of the bottom 5 nodes) indicated the presence of three additive QTL accounting for approximately 40% of the observed phenotypic variation for parthenocarpic fruit set (Table 5). All analyses returned consistent results with QTL identified on chromosome 6 at 3.3 cM, chromosome 6 at 80.0 cM, and chromosome 7 at 24.1 - 32.1 cM (Table 5). These QTL were also consistent with those identified in the analyses of the datasets collected for experiment 1 alone and experiments 2 and 3 combined. The QTL on chromosome 5 at 32.3 - 54.7 cM which was present in all other QTL analyses was not detected in data from the first 10 nodes of plant growth (Figure 3, Figure 4). The QTL effects of the QTL on chromosome 6 at 3.3 and 80.0 cM, respectively, are nearly identical (Figure 3, Figure 4, Table 5). The QTL on chromosome 7 at 24.1 - 32.1 cM is the QTL of strongest effect in data from the first 10 nodes of plant growth (Figure 3, Figure 4, Table 5).

Favorable alleles at each QTL were the same as found in experiments 1-3 for data collected from the first 30 nodes. MIM did not detect epistatic interactions between QTL. None of the analyses detected QTL with significant dominance effects at alpha = 0.05.

The strong effect of the QTL from chromosome 7 in data collected from the first 10 nodes of plant growth suggests that it may be important in very early parthenocarpic fruit set. If true, this potentially explains the inability to detect this QTL in the QTL analysis of experiment 1, as early fruit set was disrupted by stress related to plant crowding. Overall, phenotypic selection for parthenocarpic fruit set is possible with as few as 10 nodes of plant growth (Chapter 1 Figure 7, Table 5). However, there is some risk of omitting QTL that may be important to fully maximizing parthenocarpic potential, such as the QTL on chromosome 5. It is proposed here that phenotypic selection should be done with 20 nodes of plant growth as active fruit set of the first flush of fruits is often continuing at node 10 and beyond in the parthenocarpic parental line '2A' (Chapter 1 Figure 6). Limiting evaluation to 10 nodes (only 5 scorable nodes) may be too strict and plants should be allowed to finish set of the first flush of fruits (four - seven fruits for plants with high parthenocarpic potential) to maximize observed expression.

The new approach employed by this study for accurate phenotyping of parthenocarpic fruit set in cucumber by focusing on early fruit initiation and development was highly effective. QTL analyses of pooled data from experiments 1-3 revealed seven additive QTL accounting for 73.0% (CIM) - 75.5% (MIM) of the observed phenotypic variation. We propose that these QTL be designated as *parth2.1* (chromosome 2 at 0.0 – 0.9 cM), *parth4.1* (chromosome 4 at 83.2 – 87.7 cM), *parth5.1* (chromosome 5 at 32.3 - 54.7 cM), *parth6.1* (chromosome 6 at 0.0 - 9.7 cM), *parth6.2* (chromosome 6 at 80.0 - 83.0 cM), *parth6.3* (chromosome 6 at 53.0 cM), and *parth7.1* (chromosome 7 at 21.8 - 32.1 cM). Further, analyses of individual datasets obtained from

experiment 1 alone, experiments 2 and 3 combined, and the first 10 and 20 nodes of plant growth consistently indicated the presence of four QTL (*parth5.1*, *parth6.1*, *parth6.2*, and *parth 7.1*) with moderate to large effect (approximately 10 - 20%) for parthenocarpic fruit set potential. The favorable alleles at each of these four QTL are attributed to the parthenocarpic parental line '2A' with the exception of *parth6.2*, were the favorable allele is contributed by the non-parthenocarpic parental line 'Gy8'. The remaining three minor QTL (*parth2.1*, *parth4.1*, and *parth6.3*), which were not consistently found in all analyses, all had favorable alleles attributable to 'Gy8' at these loci. It seems plausible that these three minor QTL may be related to parthenocarpic fruit set and/or yield in high stress environments. Regardless, future focus on understanding the mechanism of parthenocarpic fruit development in cucumber should focus on the consensus four moderate to large effect QTL.

Comparison to Previously Identified QTL for Parthenocarpic Expression in Cucumber

A comparison of QTL associated with parthenocarpic expression in cucumber identified by this study and the one conducted by Sun et al. (2006), with another 2A×Gy8 $F_{2:3}$ population, revealed both agreement and disagreement. In the Sun et al. study, plants were grown in isolated outdoor field plots and parthenocarpic potential was measured as the number of fruit exceeding 2.8 cm in diameter during a single harvest performed when 15 % of fruit were at least 5 cm in diameter. QTL analyses were performed with MIM, CIM, and IM QTL detection approaches. Through the use of common SSR markers described in Chapter 2, a rough comparison of QTL locations can be made (Chapter 2 Addendum 2). Both studies identified three QTL on chromosome 6 at similar chromosome positions (Addendum 2). Similarly, each of these QTL were estimated to account for 10-15% of the phenotypic variation. None of the other QTL

identified by either study could be validated by both studies. The Sun et al. study also concluded there were four major genomic regions associated with parthenocarpic expression, although not the same four regions identified by this study.

The Sun et al. study suggested that some QTL associated with parthenocarpic expression corresponded with QTL identified by Fazio et al. (2003) for fruit number per plant at first harvest (yield). To investigate this observation, an additional comparison was made between the QTL identified by this study and QTL associated with fruit yield by Fazio et al. (Addendum 2). A strong association between the QTL for the two traits was not observed. Only the QTL on chromosome 6 at 53.0 cM (*parth6.3*) appears to overlap between the two traits. This observation adds confidence for the effectiveness of the approach taken by this study to accurately evaluate and phenotype parthenocarpic expression with minimal interference from yield as a confounding trait.

Identification of QTL Associated With Seed Size and Seed Weight

In addition to parthenocarpic fruit set, seed size and seed weight traits were observed to be segregating in this 2A×Gy8 population. A single dataset for each trait was constructed from the measurement of seed obtained from the self-pollination of F_2 plants. Each dataset was analyzed with MIM, CIM, and IM QTL detection approaches in the same manner as outlined for parthenocarpic fruit set. Analyses of data collected for seed size indicated the presence of as many as four (MIM) additive QTL accounting for 19% (CIM) - 36% (MIM) of the observed phenotypic variation (Figure 5, Figure 6, Table 6). There was minor disagreement between the QTL detection approaches on the number of QTL, with CIM and IM detecting two QTL and MIM detecting four (Table 6). However, all three approaches concurred on the presence of QTL

on chromosome 5 at 14.9 cM and chromosome 6 at 80.0 cM. Each QTL was of small to moderate effect and accounted for approximately 10% of the observed phenotypic variation. MIM detected an epistatic interaction between the QTL located on chromosomes 3 and 4 (Table 6). This is likely the reason for the discrepancy between MIM and the other QTL detection approaches as these loci are not significant when considered alone. None of the analyses detected QTL with significant dominance effects at alpha = 0.05. Favorable alleles for smaller seed size were associated with parental line '2A' at each QTL except for the QTL on chromosome 5 at 14.9 cM where favorable alleles where contributed by 'Gy8' (Table 6). The QTL of greatest interest is the QTL on chromosome 6 at 80.0 cM which corresponds to *parth6.2*.

Analyses of data collected for seed weight indicated the presence of as many as four (MIM) additive QTL accounting for 23% (CIM) - 37% (MIM) of the observed phenotypic variation (Figure 6, Figure 7, Table 7). Similar to the QTL analyses of seed size, all QTL analyses of seed weight identified QTL located on chromosome 5 at 20.4 cM and chromosome 6 at 80.0 cM (Figure 5, Figure 7, Table 6, Table 7). This is not surprising since the seed weight and seed size traits are expected to be highly correlated and this observation adds support for these two loci as true QTL (Addendum 9). Each QTL accounted for approximately 10% of the observed phenotypic variation. None of the analyses detected QTL with significant dominance effects at alpha = 0.05. Again, all favorable alleles for smaller seed size were associated with parental line '2A' at each QTL except for the QTL on chromosome 5 at 14.9 cM where favorable alleles where contributed by 'Gy8' (Table 7). MIM identified the presence of two additional QTL on chromosome 4 at 6.4 cM and chromosome 7 at 23.0 cM, which were not detected with CIM or IM (Table 7). MIM also identified an epistatic interaction between the QTL located on

chromosome 6 at 80.0 cM and chromosome 7 at 23.0 cM. This is interesting to note as both of these QTL were also identified as important to parthenocarpic fruit set.

The seed size and weight traits are particularly interesting in this population since the QTL located on chromosome 6 at 80.0 cM is important to the expression of parthenocarpic fruit set and both seed traits (Figure 8). The QTL on chromosome 6 at 80.0 cM could potentially have pleiotropic effects on both traits. The QTL for the seed traits on chromosome 5 at 14.9 - 20.4 cM does not appear to be the same as *parth5.1* (Figure 8). However, the LOD curves for each trait do intersect above the LOD threshold of alpha = 0.10 at approximately 25.7 cM on chromosome 5. Given the low marker density, it remains a low possibility that these two QTL could be identifying the same locus on chromosome 5. The QTL associated with seed size and weight on chromosome 4 at 6.4 - 10.4 cM and chromosome 5 at 14.9 - 20.4 cM are confirmed by a more in depth study of these traits by Wang et al. (2014). However, the QTL on chromosome 6 at 80.0 cM was not identified by Wang et al. and supports this locus as a unique locus affecting seed size in this population.

Identification of Candidate Genes

After the identification of four consistent QTL for the inheritance of parthenocarpic fruit set, the genomic sequence neighboring each QTL was explored for candidate genes by BLASTn search of Arabidopsis mRNA sequences. The functions of all gene homolog matches were investigated and evaluated for potential to influence parthenocarpic fruit set. However, since the consensus in the available literature implicates parthenocarpic expression as being under plant hormonal control, gene homologs with function in hormonal pathways were closely evaluated as potential candidates. Each QTL region included several plant hormone related gene homologs. The

parth6.2 QTL possessed the narrowest 1.5 LOD interval so all candidate genes in this region were first considered as the basis of a potential genetic mechanism (Addendum 3, Addendum 4). One gene homolog in this region was the brassinosteroid receptor BRI1 (Addendum 4). Interestingly, a previous study by Fu et al. (2008) found the exogenous application of synthetic brassinosteroids to induce parthenocarpic expression in a non-parthenocarpic cucumber cultivar. In addition, the application of brassinazole, a brassinosteroid biosynthesis inhibitor, inhibited parthenocarpic expression in a parthenocarpic cucumber cultivar (Fu et al., 2008). Another piece of evidence comes from the observations of decreased seed size in plants deficient in brassinosteroid signaling and BRI1 mutants (Huang et al., 2012; Morinaka et al., 2006; Nakagawa et al., 2012; Tanabe et al., 2005). Since the QTL on chromosome 6 at 80.0 - 83.0 cM (*parth6.2*) was shared for both parthenocarpic fruit set and seed size, BRI1 is a promising preliminary candidate gene for this QTL.

With BRI1 identified as a candidate gene at *parth6.2*, the other three QTL regions were explored for gene homologs that may potentially interact with BRI1. Two homologs of BAK1 were found to be located in the QTL regions of *parth6.1* and *parth7.1* (Addendum 3, Addendum 5, Addendum 6). Although BAK1, another LRR-RLK, is unable to perceive brassinosteroids, it acts a co-receptor to BRI1 and enhances BRI1 activity (Kim and Wang, 2010; Li et al, 2002; Nam and Li, 2002, Russinova et al., 2004, Wang et al., 2008). As demonstrated in both Arabidopsis and tomato, the BAK1 protein also directly interacts with the BRI1 protein and together they form a heterodimer (Bajwa et al., 2013; Russinova et al., 2004; Wang et al., 2008). Since a direct interaction between these two proteins has been shown to occur and along with the fact that homologs of these interacting proteins were located within three of the four QTL regions identified for parthenocarpic fruit set, this occurrence may be more than coincidental.

Another potential candidate gene, phosphatase 2A B' alpha (PP2A), was identified in the QTL region of *parth6.1*. PP2A has been demonstrated to be an important component of the brassinosteroid signaling pathway (Tang et al., 2011). PP2A promotes brassinosteroid signaling by dephosphorylation and consequent activation of BZR1 (Tang et al., 2011). PP2A B' alpha mutants obtained through T-DNA insertions displayed phenotypes similar to slight BRI1 mutant phenotypes. Both BAK1 and PP2A are promising candidate genes for parthenocarpic fruit set in cucumber worthy of further investigation. However, the focus of this study choose to pursue BAK1 due to the fact that BAK1 homologs were located within two of the four QTL identified for parthenocarpic fruit set and this occurrence was considered to be more than coincidental. In addition, potential mutations to either BAK1 or PP2A would likely result in similarly lower levels of activated BZR1.

A thorough examination of gene homolog matches in the fourth QTL region, *parth5.1*, identified the DELLA proteins Gibberellic Acid Insensitive (GAI) and Repressor of GA1-3 (RGA) as potential candidate genes (Addendum 3, Addendum 7). DELLAs are transcription regulators that restrict plant growth and negatively regulate gibberellin growth responses (Dill and Sun, 2001; Dill et al., 2004; Li et al., 2012; Sun, 2011). Silencing or loss of DELLA proteins has been found to induce facultative parthenocarpic expression (Carrera et al., 2012; Dorcey et al., 2009; Fuentes et al., 2012; Marti et al., 2007). Microarray studies have observed significant overlap in the genes affected in the brassinosteroid insensitive mutant *bri-116* and the gibberellic acid insensitive mutant *ga1-3* (rga), suggesting both have similar effects on a large number of common genes (Bai et al., 2012; Cheminant et al., 2011; Sun et al., 2010). DELLA proteins RGA and GAI have also been found to directly interact with BZR1 by binding to the active dephosphorylated form of BZR1 and inhibiting its transcriptional activity (Bai et al., 2012;

Gallego-Bartolome et al., 2012; Li et al., 2012). This demonstrates the role of DELLAs in negatively regulating the brassinosteroid pathway. The presence of a gene homolog found to interact with a downstream product of BRI1 in the fourth QTL region supported the proposal of a possible crosstalk mechanism between the gibberellin and brassinosteroid signaling pathways.

A Potential Mechanism for Parthenocarpic Fruit Set

Our findings point to a potential mechanism for parthenocarpic fruit set based on crosstalk between the brassinosteroid and gibberellin signaling pathways. Parthenocarpic expression can be viewed as the release of fruit growth inhibition without pollination. Following this theme, parthenocarpy has been found to be under control of GA signaling, which in part includes removal of growth inhibition imposed by DELLAs through GA induced DELLA degradation (Dorcey et al., 2009; Fuentes et al. 2012; Marti et al., 2007; Serrani et al., 2008; Serrani et al., 2010). Unfortunately, limited evidence exists in the literature connecting brassinosteroids to parthenocarpic expression. Further, none of the typical phenotypic responses observed in brassinosteroid deficient and insensitive plants such as: dwarfism, dark green leaves, altered leaf and vascular morphology, delayed senescence and flowering, and male infertility were observed in the 2A×Gy8 population (Altmann, 1999; Clouse et al., 1996; Li and Chory, 1997; Noguchi et al., 1999; Yamamuro et al., 2000; Montoya et al., 2002). However, preliminary observations made prior to commencing this experiment agree with the observations of Sun et al. (2006) in noting that ‘2A’ does have reduced plant vigor in comparison to ‘Gy8’. Although not conclusive, this may be an indication of reduced brassinosteroid perception or biosynthesis. Indeed, not all mutant alleles of BRI1 result in severe phenotypes and some may closely resemble the wildtype phenotype (Morinaka et al., 2006; Noguchi et al., 1999). BRI1 null

mutants have been found to accumulate very high levels of brassinosteroids; while partial loss of function alleles have also been found to have elevated levels (Bancos et al., 2002; Noguchi et al., 1999). This likely is a result of the inability to perceive brassinosteroids at the receptor and consequently a brassinosteroid dependent biosynthesis feedback mechanism fails to activate (Bancos et al., 2002; Mathur et al., 1998). The effect of brassinosteroids on endogenous gibberellin levels is still unresolved. Exogenous application of brassinosteroids has been found to induce the expression of genes involved in gibberellin biosynthesis in brassinosteroid deficient and wildtype plants (Bouquin et al., 2001; Li et al., 2012; Wang et al., 2009). However, this effect was not observed in brassinosteroid insensitive plants (Bouquin et al., 2001). Conversely, measurements of bioactive gibberellic acid and its precursors revealed brassinosteroid deficient and insensitive mutants produced significantly elevated levels of gibberellic acid precursors (Jager et al., 2005; Nadhzimov et al., 1988). In pea, examination of both brassinosteroid deficient and insensitive mutants revealed elevated levels of bioactive gibberellic acid, although brassinosteroid deficient mutants were found not to be statistically different from wildtype plants despite an observed 2.7 fold increase (Jager et al., 2005). Further, the loss of DELLA protein function may also promote gibberellin signaling response through loss of inhibition (Harberd et al., 2009; Sun, 2011; Weston et al., 2008). These observations along with the candidate genes identified by this study imply that the gibberellin signaling pathway may potentially be a core component of parthenocarpic fruit set in cucumber.

Construction of a proposed mechanism for parthenocarpic fruit set begins with BRI1. Reports of decreased seed size in BRI1 mutants and induction of parthenocarpic expression with the application of exogenous brassinosteroids support BRI1 as a candidate gene in this population. Since the parthenocarpic parental line ‘2A’ has a small seed size, it must contain a

BRI1 allele with at least partial loss of function or expression. The BRI1 co-receptor, BAK1, has been identified as a candidate gene at two of the four QTL. If parthenocarpic fruit set in cucumber is partially the result of a loss in brassinosteroid perception, '2A' presumably contains alleles of BAK1 with at least partial loss of function or expression at both QTL. Diminished efficacy in binding between BRI1 and BAK1 proteins would further decrease brassinosteroid perception. With diminished perception of brassinosteroids, endogenous brassinosteroid levels accumulate and could enhance gibberellin biosynthesis.

Due to the lack of typical phenotypic responses observed with BRI1 defective mutants, a complete loss of function at the BRI1 locus in this population is unlikely. The BRI1 mutant observed in this population likely represents a partial loss of function allele of BRI1 that appears phenotypically similar to wildtype plants as has been observed in Arabidopsis and rice (Morinaka et al., 2006; Noguchi et al., 1999). A partial loss of function allele of BRI1 would allow for a low level of brassinosteroid signaling through homodimerization in the absence of functioning BAK1 proteins (Wang et al., 2008). Alternatively, the point of mutation may not occur in the structure of BRI1 itself but may occur at transcription recognition sites altering BRI1 expression. Morinaka et al. (2006) demonstrated that transgenic suppression of BRI1 expression could produce very mild non-dwarf brassinosteroid related phenotypes.

The identification of a DELLA protein with homology to GAI and RGA in the fourth QTL region fits as a possible candidate gene if parthenocarpic fruit set is a partial result of loss or degradation of DELLA proteins. Loss of inhibition due to DELLA proteins enhances plant responses to gibberellins (Harberd et al., 2009; Sun, 2011; Weston et al., 2008). In this proposed model, '2A' contains a defective DELLA protein that leads to enhanced response to gibberellin signaling. The decrease in brassinosteroid perception serves to further enhance this response

through alteration of gibberellin biosynthesis. A mechanism revolving around direct crosstalk between DELLAs and BZR1 of the brassinosteroid signaling pathway as reported by Bai et al. (2012), Gallego-Bartolome et al. (2012), and Li et al. (2012), is not a likely mechanism for this interaction as BZR1 levels would also be expected to be decreased in BRI1 mutants. Further, it has also been demonstrated that BRI1 mutants do not affect the expression levels of DELLA proteins (Li et al., 2012). Alternatively, we propose an interaction between the increase in endogenous brassinosteroid levels and gibberellin biosynthesis as promoting parthenocarpic fruit set. A similar model has been suggested for studies of auxin and cytokinin induced parthenocarpic expression where both auxin and cytokinin were found to promote parthenocarpic fruit set through enhanced biosynthesis of gibberellins (Ding et al., 2013; Fuentes et al., 2012; Serrani et al., 2008; Serrani et al., 2010; Weiss and Ori, 2007). Due to the complexity of hormone crosstalk it should be expected that increases in endogenous brassinosteroid levels may also directly or indirectly affect the signaling pathways of other hormones involved in parthenocarpic fruit set (i.e. auxin).

Montoya et al. (2005) observed strong expression of brassinosteroid C-6 oxidase, a gene involved in brassinosteroid biosynthesis, in the carpels of developing flowers and associated with seed development in developing tomato fruits. During fruit development, the strongest expression was observed during early seed development. Further, grafting experiments revealed that brassinosteroids were not transported from the site of synthesis. A lack of endogenous brassinosteroid transport has also been reported by others (Bishop and Yokota, 2001; Symons and Reid, 2004). Organ specific expression of brassinosteroid biosynthetic genes has also been observed with the highest levels of expression observed in pollen, seeds, and fruits (Bajguz and Tretyn, 2003; Bancos et al., 2002; Montoya et al., 2005; Shimada et al., 2003; Symons et al.,

2006). These observations suggest that the proposed model could selectively affect flowering and early fruit development without significantly affecting other plant organs. In addition, elevated brassinosteroid levels induced by defective brassinosteroid perception at the BRI1/BAK1 complex may also mimic levels observed during seed development and potentially promote parthenocarpic expression through the proposed model.

The biggest challenge to this proposed model is the fact that the favorable allele at *parth6.2* is contributed by the non-parthenocarpic parental line 'Gy8' (Table 1, Table 4, Table 5). If the proposed model is true, this means that the wildtype BRI1 allele is favorable in combination with null or partial loss of function DELLA and BAK1 alleles for parthenocarpic fruit set. However, this observation does not eliminate the proposed model. The wildtype BRI1 allele in combination with null or partial loss of function BAK1 alleles would still exhibit weakened brassinosteroid perception due to the inability to form the BRI1/BAK1 heterodimer complex. As noted, BRI1 can homodimerize and initiate basal brassinosteroid signaling responses, allowing for activation of growth promoting brassinosteroid response genes (Wang et al., 2005; Wang et al., 2008). This may not only serve to alleviate the deleterious effects of severely diminished brassinosteroid perception, but also contribute to plant fitness enabling increased fruit set. Nearly all F_3 families with the highest potential for parthenocarpic fruit set were homozygous for the 'Gy8' allele at *parth6.2*. However, only 12 F_3 families achieved higher measurements of parthenocarpic fruit set than the parthenocarpic parental line '2A' (Addendum 8). One possible reason why so few lines were found to exceed '2A' may be related to the decision to collect phenotypic data on parthenocarpic fruit set at a single time when plants had reached 35 nodes in growth. As discussed in Chapter 1, it is likely that '2A' and F_3 families and with high potential for parthenocarpic fruit set were never observed at their full potential due

to the confounding trait of an individual plants capacity for fruit load. In this case, the ideal genotype for parthenocarpic fruit set maybe one that includes the BRI1 allele of 'Gy8', but a noticeable increase in parthenocarpic expression is not observed over '2A' because the fruit load capacity of an individual plant has already been exceeded. An ideal genotype including the wildtype BRI1 allele may also explain why some F_3 families with high parthenocarpic potential did not have small seeds (Addendum 8, Addendum 9). What remains to be answered is if the favorable allele possessed by 'Gy8' is favorable in all gene combinations. Previous studies reporting linkage between the *F* locus (gynoecy) and parthenocarpy in cucumber support the existence of a major QTL in proximity to *parth6.2* as identified in this study (de Ponti and Garretsen, 1976) (Addendum 2). However, if the favorable allele at this locus were contributed by the non-parthenocarpic 'Gy8', it would be in contradiction to those previous studies. Further, the question of why the highly parthenocarpic line '2A' is capable of high parthenocarpic potential while lacking the wildtype BRI1 allele suggests that the ideal genotype at the BRI1 locus may be dependent on the genotypes at the other candidate gene loci, and in particular BAK1. The fact that the favorable allele at the *parth6.2* QTL is contributed from 'Gy8' in this population may also support BRI1 as the candidate gene by suggesting the elimination of other candidate genes. For example, geranylgeranyl pyrophosphate synthase (GGPS), encodes a precursor to gibberellin biosynthesis and also appears in the BLASTn search for this region, would not be a good fit when considering the favorable allele for this candidate gene is contributed by the non-parthenocarpic parent 'Gy8' (Kuntz et al. 1992). The assumption if GGPS were the candidate gene would be that increased gibberellin biosynthesis would induce parthenocarpic expression. However, since the favorable allele for parthenocarpic fruit set is contributed by the non-parthenocarpic parent 'Gy8', this assumption would not fit.

Alignment of Re-sequencing Data for BRI1 and BAK1 Genes

With a proposed genetic mechanism for the inheritance of parthenocarpic fruit set in cucumber, an attempt to identify casual polymorphisms in the cucumber homologs of BRI1 and BAK1 was made. The gene sequence and structure including transcription initiation sites, exons, and introns for BRI1 and BAK1 were extracted from the Gy14 draft genome assembly with the FGENESH utility provided by Softberry. Only the copy of BAK1 found at *parth6.1* was investigated. Using the gene sequences from 'Gy14' as a reference, whole genome re-sequencing data for '2A' and 'Gy8' (Chapter 2) were compared by sequence alignment. Numerous single nucleotide polymorphisms were identified between the parental lines for each gene. To identify any potential codon changes that could be attributed to nucleotide polymorphisms, the assembled gene sequences for each parent were imported into the FGENESH utility to identify predicted protein sequences (Addendum 10, Addendum 11). Interestingly, the predicted protein sequences for each gene revealed a single amino acid change between the parental lines (Addendum 10, Addendum 11).

Sanger sequencing of each gene in the parental lines was performed in order to validate the polymorphisms identified with the re-sequencing data. None of the nucleotide polymorphisms were confirmed as true polymorphisms with Sanger sequencing. The reliability of the re-sequencing data was already questioned by the numerous false polymorphisms identified during attempts to fill gaps in the genetic linkage map in Chapter 2. To further validate, dCAPS markers were designed for both of the nucleotide polymorphisms predicted to result in a codon change from the re-sequencing data. At least four additional dCAPS markers were designed for nucleotide polymorphisms predicted to lie in the intron regions for each gene. None of the dCAPS markers were found to be polymorphic between the parental lines. The

failure to identify polymorphisms in the gene sequences of BRI1 and BAK1 by this study does not mean that polymorphisms do not exist in either candidate gene. With the quality of the re-sequencing data in question, each gene should ideally be sequenced fully by Sanger sequencing. In addition, more of the surrounding sequence should be investigated to allow for errors in gene prediction and to also allow for polymorphisms that effect transcription.

Future Focus

The QTL identified for parthenocarpic fruit set by this study are valuable to cucumber breeders interested in developing parthenocarpic cultivars and to researchers interested in the inheritance and mechanism of parthenocarpic fruit set. However, future efforts will be needed in fine mapping the QTL regions identified here in order to either confirm the proposed candidate genes or identify new ones. As seen in the available literature, manipulation of most plant hormones or hormone transport mechanisms can result in parthenocarpic expression. Because of this, the candidate genes identified here must also be validated in other parthenocarpic cucumber populations to explore whether a single or multiple sources of parthenocarpy exist. Finally, the mechanism proposed here warrants further investigation but future studies must still consider other candidate genes identified from the QTL regions.

Literature Cited

Altmann, T. 1999. Molecular physiology of brassinosteroids revealed by the analysis of mutants. Planta. 208:1-11.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Bai, M.Y., J.X. Shang, E. Oh, M. Fan, Y. Bai, R. Zentella, T.P. Sun, and Z.Y. Wang. 2012. Brassinosteroid, gibberellin and phytochrome impinge on a common transcription module in *Arabidopsis*. Nat. Cell Biol. 14:810-817.

Bajguz, A. and A. Tretyn. 2003. The chemical characteristic and distribution of brassinosteroids in plants. Phytochemistry. 62:1027-1046.

Bancos, S., T. Nomura, T. Sato, G. Molnar, G.J. Bishop, C. Koncz, T. Yokota, F. Nagy, and M. Szekeres. 2002. Regulation of transcript levels of the Arabidopsis cytochrome P450 genes involved in brassinosteroid biosynthesis. Plant Physiol. 130:504-513.

Bajwa, V.S., X. Wang, R.K. Blackburn, M.B. Goshe, S.K. Mitra, E.L. Williams, G.J. Bishop, S. Krasnyanski, G. Allen, S.C. Huber, and S.D. Clouse. 2013. Identification and functional analysis of tomato BRI1 and BAK1 receptor kinase phosphorylation sites. Plant Physiol. 163:30-42.

Beraldi, D., M.E. Picarella, G.P. Soressi, and A. Mazzucato. 2004. Fine mapping of the parthenocarpic fruit (*pat*) mutation in tomato. Theor. Appl. Genet. 108:209-216.

Bishop, G.J. and T. Yokota. 2001. Plant steroid hormones, brassinosteroid: current highlights of molecular aspects on their synthesis/metabolism, transport, perception and response. Plant Cell Physiol. 42:114-120.

Bouquin, T., C. Meier, R. Foster, M.E. Nielsen, and J. Mundy. 2001. Control of specific gene expression by gibberellin and brassinosteroid. Plant Physiol. 127:450-458.

Broman, K.W., H. Wu, S. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 19:889-890.

Broman, K.W. and S. Sen. 2009. A Guide to QTL Mapping with R/qtl. Springer Science and Business Media. New York, NY.

Carrera, E., O. Ruiz-Rivero, L.E.P. Peres, A. Atares, and J.L. Garcia-Martinez. 2012. Characterization of the *procera* tomato mutant shows novel functions of the SIDELLA protein in the control of flower morphology, cell division and expansion, and the auxin-signaling pathway during fruit-set and development. Plant Physiol. 160:1581-1596.

Cheminant, S., M. Wild, F. Bouvier, S. Pelletier, J.P. Renou, M. Ethardt, S. Hayes, M.J. Terry, P. Genschik, and P. Achard. 2011. DELLAs regulate chlorophyll and carotenoid biosynthesis to prevent photooxidative damage during seedling deetiolation in Arabidopsis. Plant Cell. 23:1849-1860.

Clouse, S.D., M. Langford, and T.C. McMorris. 1996. A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. Plant Physiol. 111:671-678.

Clouse, S.D. 2011. Brassinosteroid signal transduction: from receptor kinase activation to transcriptional networks regulating plant development. Plant Cell. 23:1219-1230.

Dill, A. and T.P. Sun. 2001. Synergistic derepression of gibberellin signaling by removing RGA and GAI function in *Arabidopsis thaliana*. Genetics. 159:777-785.

Dill, A., S.G. Thomas, J. Hu, C.M. Steber, and T.P. Sun. 2004. The *Arabidopsis* F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. Plant Cell. 16:1392-1405.

Ding, J., B. Chen, X. Xia, W. Mao, K. Shi, Y. Zhou, and J. Yu. 2013. Cytokinin-induced parthenocarpic fruit development in tomato is partly dependent on enhanced gibberellin and auxin biosynthesis. PLoS One. 8:e70080.

Divi, U.K. and P. Krishna. 2009. Brassinosteroid: a biotechnological target for enhancing crop yield and stress tolerance. New Biotechnol. 26:131-136.

Dorcey, E., C. Urbez, M.A. Blazquez, J. Carbonell, and M.A. Perez-Amador. 2009. Fertilization-dependent auxin response in ovules triggers fruit development through the modulation of gibberellin metabolism in *Arabidopsis*. Plant J. 58:318-332.

El-Shawaf, I.I.S. and L.R. Baker. 1981. Inheritance of parthenocarpic yield in gynoecious pickling cucumber for once-over mechanical harvest by diallel analysis of six gynoecious lines. J. Amer. Soc. Hort. Sci. 106:365-370.

Fazio, G., J.E. Staub, and M.R. Stevens. 2003. Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor. Appl. Genet. 107:864-874.

Fos, M., F. Nuez, and J.L. Garcıa-Martınez. 2000. The gene pat-2, which induces natural parthenocarpy, alters the gibberellin content in unpollinated tomato ovaries. Plant Physiol. 122:471-480.

Fu, F.Q., W.H. Mao, Y.H. Zhou, T. Asami, and J.Q. Yu. 2008. A role of brassinosteroids in early fruit development in cucumber. J. Exp. Bot. 59:2299-2308.

Fuentes, S., K. Ljung, K. Sorefan, E. Alvey, N.P. Harberd, and L. Ostergaard. 2012. Fruit growth in *Arabidopsis* occurs via DELLA-dependent and DELLA-independent gibberellin responses. Plant Cell. 24:3982-3996.

Gallego-Bartolome, J., E.G. Minguet, F. Grau-Enguix, M. Abbas, A. Locascio, S.G. Thomas, D. Alabadi, and M.A. Blazquez. 2012. Molecular mechanism for the interaction between gibberellin and brassinosteroid signaling pathways in *Arabidopsis*. Proc. Natl. Acad. Sci. 109:13446-13451.

Gillaspy, G., H. Ben-David, and W. Gruissem. 1993. Fruits: A developmental perspective. Plant Cell. 5:1439-1451.

Gorguet, B., P.M. Eggink, J. Ocana, A. Tiwari, D. Schipper, R. Finkers, R.G.F. Visser, and A.W. van Heusden. 2008. Mapping and characterization of novel parthenocarpy QTLs in tomato. Theor. Appl. Genet. 116:755-767.

Harberd, N.P., E. Belfield, and Y. Yasumura. 2009. The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: how and "inhibitor of an inhibitor" enables flexible response to fluctuating environments. Plant Cell. 21:1328-1339.

Hawthorn, L.R. and R. Wellington. 1930. Geneva, a greenhouse cucumber that develops fruit without pollination. NY (Geneva) Agric Exp. Stat. Bull. 580:223-234.

He, J.X., J.M. Gendron, Y. Sun, S.S. Gampala, N. Gendron, C.Q. Sun, and Z.Y. Wang. 2005. BZR1 is a transcriptional repressor with dual roles in brassinosteroid homeostasis and growth responses. Science. 307:1634-1638.

Huang, H., W. Jiang, Y. Hu, P. Wu, J. Zhu, W. Liang, Z.Y. Wang, and W. Lin. 2012. BR signal influences *Arabidopsis* ovule and seed number through regulating related genes expression by BZR1. Mol. Plant. 6:456-469.

Jager, C.E., G.M. Symons, J.J. Ross, J.J. Smith, and J.B. Reid. 2005. The brassinosteroid growth response in pea is not mediated by changes in gibberellin content. Planta. 221:141-148.

Kim T.W. and Z.Y. Wang. 2010. Brassinosteroid signal transduction from receptor kinases to transcription factors. Annu. Rev. Plant Biol. 61:681-704.

Kuntz, M., S. Romer, C. Suire, P. Huguney. J.H. Weil, R. Schantz, and B. Camara. 1992. Identification of a cDNA for the plastid-located geranylgeranyl pyrophosphate synthase from Capsicum annuum: correlative increase in enzyme activity and transcript level during fruit ripening. Plant J. 2:25-34.

Kvasnikov, B.V., N.T. Rogova, S.I. Tarakonova, and I. Ignatova. 1970. Methods of breeding vegetable crops under the covered ground. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 42:45-57.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, and D.G. Higgins. 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23:2947-2948.

Li, J. and J. Chory. 1997. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. Cell. 90:929-938.

Li, J., J. Wen, K.A. Lease, J.T. Doke, F.E. Tax, and J.C. Walker. 2002. BAK1, an *Arabidopsis* LRR receptor-like kinase, interacts with BRI1 and modulates brassinosteroid signaling. Cell. 110:213-222.

Li, Q.F., C. Wang, L. Jiang, S. Li, S.S.M. Sun, and J.X. He. 2012. An interaction between BZR1

and DELLAs mediates direct signaling crosstalk between brassinosteroids and gibberellins in *Arabidopsis*. Sci. Signal. 5:ra72.

Manichaikul, A., J.Y. Moon, S. Sen, B.S. Yandell, and K.W. Broman. 2009. A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. *Genetics*. **181**:1077-1086.

Marti, C., D. Orzaez, P. Ellul, V. Moreno, J. Carbonell, and A. Granell. 2007. Silencing of *DELLA* induces facultative parthenocarpy in tomato fruits. Plant J. 52:865-876.

Mather, J., G. Molnar, S. Fujioka, S. Takatsuto, A. Sakurai, T. Yokota, G. Adam, B. Voigt, F. Nagy, C. Maas, J. Schell, C. Koncz, and M. Szekeres. 1998. Transcription of the *Arabidopsis CPD* gene, encoding a steroidogenic cytochrome P450, is negatively controlled by brassinosteroids. Plant J. 14:593-602.

de Menezes, C.B., W.R. Maluf, S.M. de Azevedo, M.V. Faria, I.R. Nascimento, D.W. Nogueira, L.A.A. Gomes, and E. Bearzoti. 2005. Inheritance of parthenocarpy in summer squash (*Cucurbita pepo* L.) Genet. Mol. Res. 4:39-46.

Meshcherov, E.T. and L.W. Juldasheva. 1974. Parthenocarpy in cucumber. Trudy-po-Prikladnoi-Botanike-Genetiki-I-Selektsii. 51:204-213.

Miyatake, K., T. Saito, S. Negoro, H. Yamaguchi, T. Nunome, A. Ohyama, and H. Fukuoka. 2012. Development of selective markers linked to a major QTL for parthenocarpy in eggplant (*Solanum melongena* L.). Theor. Appl. Genet. 124:1403-1413.

Montoya, T., T. Nomura, K. Farrar, T. Kaneta, T. Yokota, and G.J Bishop. 2002. Cloning the tomato *curl3* gene highlights the putative dual role of the leucine-rich repeat receptor kinase tBRI1/SR160 in plant steroid hormone and peptide hormone signaling. Plant Cell. 14:3163-3176.

Montoya, T., T. Nomura, T. Yokota, K. Farrar, K. Harrison, J.G.D. Jones, T. Kaneta, Y. Kamiya, M. Szekeres, and G.J. Bishop. 2005. Patterns of *Dwarf* expression and brassinosteroid accumulation in tomato reveal the importance of brassinosteroid synthesis during fruit development. Plant J. 42:262-269.

Morinaka, Y., T. Sakamoto, Y. Inukai, M. Agetsuma, H. Kitano, M. Ashikari, and M.Matsuoka. 2006. Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. Plant Physiol. 141:924-931.

Nadhzimov, U.K., S.C. Jupe, M.G. Jones, and I.M. Scott. 1988. Growth and gibberellin relations of the extreme dwarf d^x tomato mutant. Physiol. Plant. 73:252-256.

Nam K.H. and J. Li. 2002. BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. Cell. 110:203-212.

Nakagawa, H., A. Tanaka, T. Tanabata, M. Ohtake, S. Fujioka, H. Nakamura, H. Ichikawa, and M. Mori. *SHORT GRAIN1* decreases organ elongation and brassinosteroid response in rice. Plant Physiol. 158:1208-1219.

Noguchi, T., S. Fujioka, S. Choe, S. Takatsuto, S. Yoshida, H. Yuan, K. Feldmann, and F.E. Tax. 1999. Brassinosteroid-insensitive dwarf mutants of Arabidopsis accumulate brassinosteroids. Plant Physiol. 121:743-752.

Olimpieri, I., F. Silicato, R. Caccia, L. Mariotti, N. Ceccarelli, G.P. Soressi, and A. Mazzucato. 2007. Tomato fruit set driven by pollination or by the parthenocarpic fruit allele are mediated by transcriptionally regulated gibberellin biosynthesis. Planta. 226:877-888.

Pandolfini, T. 2009. Seedless fruit production by hormonal regulation of fruit set. Nutrients. 1:168-177.

Pascual. L., J.M. Blanca, J. Canizares, and F. Nuez. 2009. Transcriptomic analysis of tomato carpel development reveals alterations in ethylene and gibberellins synthesis during *pat3/pat4* parthenocarpic fruit set. BMC Plant Biol. 9:67.

Pike, L.M. and C.E. Peterson. 1969. Inheritance of parthenocarpy in the cucumber (*Cucumis sativus* L.) Euphytica. 18:101-105.

de Ponti, O.M.B. and F. Garretsen. 1976. Inheritance of parthenocarpy in pickling cucumbers (*Cucumis sativus* L.) and linkage with other characters. Euphytica. 25:633-642.

Rozen, S. and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132:365-386.

Russinova, E., J.W. Borst, M. Kwaaitaal, A. Cano-Delgado, Y. Yin, J. Chory, and S.C. de Vries. 2004. Heterodimerization and endocytosis of *Arabidopsis* brassinosteroid receptors BRI1 and AtSERK3 (BAK1). Plant Cell. 16:3216-3229.

Serrani, J.C., M. Fos, A. Atares, and J.L. Garcia-Martinez. 2007. Gibberellin regulation of fruit-set and growth in tomato. Plant Physiol. 145:246-257.

Serrani, J.C., O. Ruiz-Rivero, M. Fos, and J.L. Garcia-Martinez. 2008. Auxin-induced fruit-set in tomato is mediated in part by gibberellins. Plant J. 56:922-934.

Serrani, J.C., E. Carrera, O. Ruiz-Rivero, L. Gallego-Giraldo, L.E.P. Peres, and J.L. Garcia-Martinez. 2010. Inhibition of auxin transport from the ovary or from the apical shoot induces parthenocarpic fruit-set in tomato mediated by gibberellins. Plant Physiol. 153:851-862.

Shimada, Y., H. Goda, A. Nakamura, S. Takatsuo, S. Fujioka, and S. Yoshida. 2003. Organ-specific expression of brassinosteroid-biosynthetic genes and distribution of endogenous brassinosteroids in Arabidopsis. Plant Physiol. 131:287-297.

Solovyev, V., P. Kosarev, I. Seledsov, and D. Vorobyev. 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 7 (Suppl 1):S10.

Srivastava, L.M. 2002. Plant Growth and Development: Hormones and Environment. Elsevier Science. San Diego, CA.

Sun, Z., J.E. Staub, S.M. Chung, and R.L. Lower. 2006. Identification and comparative analysis of quantitative trait loci (QTL) associated with parthenocarpy in processing cucumber. Plant Breeding. 125:281-287.

Sun, Y., X.Y. Fan, D.M. Cao, W. Tang, K. He, J.Y. Zhu, J.X. He, M.Y. Bai, S. Zhu, E. Oh, S. Patil, T.W. Kim, H. Ji, W.H. Wong, S.Y. Rhee, and Z.Y. Wang. 2010. Integration of brassinosteroid signal transduction with the transcription network for plant growth regulation in Arabidopsis. Dev. Cell. 19:765-777.

Sun, T.P. 2011. The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants. Curr. Biol. 21:R338-R345.

Symons, G.M. and J.B. Reid. 2004. Brassinosteroids do not undergo long-distance transport in pea. Implications for the regulation of endogenous brassinosteroid levels. Plant Physiol. 135:2196-2206.

Symons, G.M., C. Davies, Y. Shavrukov, I.B. Dry, J.B. Reid, and M.R. Thomas. 2006. Grapes on steroids: brassinosteroids are involved in grape berry ripening. Plant Physiol. 140:150-158.

Tanabe, S., M. Ashikari, S. Fujioka, S. Takatsuto, S. Yoshida, M. Yano, A. Yoshimura, H. Kitano, M. Matsuoka, Y. Fujisawa, H. Kato, and Y. Iwasaki. 2005. A novel cytochrome P450 is implicated in brassinosteroid biosynthesis via the characterization of a rice dwarf mutant, *dwarf11*, with reduced seed length. Plant Cell. 17:776-790.

Tang, W., M. Yuan, R. Wang, Y. Yang, C. Wang, J.A. Oses-Prieto, T.W. Kim, H.W. Zhou, Z. Deng, S.S. Gampala, J.M. Gendron, E.M. Jonassen, C. Lillo, A. DeLong, A.L. Burlingame, Y. Sun, and Z.Y. Wang. 2011. PP2A activates brassinosteroid-responsive gene expression and plant growth by dephosphorylating BZR1. Nat. Cell. Biol. 13:124-131.

Van Ooijen, J.W. and R.E. Voorrips. 2001. JoinMap 3.0: Software for the calculation of genetic maps. Plant Research International, Wageningen, The Netherlands.

Vardi, A., I. Levin, and N. Carmi. 2008. Induction of seedlessness in citrus: from classical techniques to emerging biotechnological approaches. J. Am. Soc. Hortic. Sci. 133:117-126.

Vivian-Smith, A. and A.M. Koltunow. 1999. Genetic analysis of growth regulator-induced parthenocarpy in *Arabidopsis*. Plant Physiol. 121:437-451.

Vriet, C., E. Russinova, and C. Reuzeau. 2012. Boosting crop yields with plant steroids. Plant Cell. 24:842-857.

Wang, X., X.Q. Li, J. Meisenhelder, T. Hunter, S. Yoshida, T. Asami, and J. Chory. 2005. Autoregulation and homodimerization are involved in the activation of the plant steroid receptor BRI1. Dev. Cell. 8:855-865.

Wang, X., U. Kota, K. He, K. Blackburn, J. Li, M.B. Goshe, S.C. Huber, and S.D. Clouse. 2008. Sequential transphosphorylation of the BRI1/BAK1 receptor kinase complex impacts early events in brassinosteroid signaling. Dev. Cell. 15:220-235.

Wang, L, Z. Wang, Y. Xu, S.H. Joo, S.K. Kim, Z. Xue, Z. Xu, Z. Wang, and K. Chong. 2009. *OsGSR1* is involved in crosstalk between gibberellins and brassinosteroids in rice. Plant J. 57:498-510.

Wang, M., H., Miao, S. Zhang, S. Liu, S. Dong, Y. Wang, and X. Gu. 2014. Inheritance analysis and QTL mapping of cucumber seed size. Acta Hortic. Sinica. 14:63-72.

Weiss, D. and N. Ori. 2007. Mechanisms of cross talk between gibberellin and other hormones. Plant Physiol. 144:1240-1246.

Weston, D.E., R.C. Elliott, D.R. Lester, C. Rameau, J.B. Reid, I.C. Murfet, and J.J. Ross. 2008. The pea DELLA proteins LA and CRY are important regulators of gibberellin synthesis and root growth. Plant Physiol. 147:199-205.

Yamamuro, C., Y. Ihara, X. Wu, T. Noguchi, S. Fujioka, S. Takatsuto, M. Ashikari, H. Kitano, and M. Matsuoka. 2000. Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and banding of the lamina joint. Plant Cell. 12:1591-1606.

Yin, Y, D. Vafeados, Y. Tao, S. Yoshida, T. Asami, and J. Chory. 2005. A new class of transcription actors mediates brassinosteroid-regulated gene expression in *Arabidopsis*. Cell. 120:249-259.

Yu, X., L. Li, J. Zola, M. Aluru, H. Ye, A. Foudree, H. Guo, S. Anderson, S. Aluru, P. Liu, S. Rodermel, and Y. Yin. 2011. A brassinosteroid transcriptional network revealed by genome-wide identification of BES1 target genes in *Arabidopsis thaliana*. Plant J. 65:634-646.

Yang, L.M., D.H. Koo, Y.H. Li, X.J. Zhang, F.S. Luan, M.J. Havey, J.M. Jiang, and Y. Weng. 2012. Chromosome rearrangements during domestication of cucumber as revealed by high-density genetic mapping and draft genome assembly. Plant J. 71:895-906.

Table 1. QTL mapping results for parthenocarpic fruit set in the pooled data obtained from experiments 1-3. The pooled data consists of 205 F_3 families represented by 11 individuals each. Phenotypic data was collected from nodes 5-30 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			62.66	75.53					
	1	2	0.0	9.13	5.60	-0.37	-0.11	0.0	0.0	7.2
	2	4	83.2	8.00	4.80	-0.34	-0.09	73.8	83.2	87.7
	3	5	32.3	12.86	8.20	0.49	0.16	26.4	32.3	52.6
	4	6	9.7	17.90	12.20	0.61	-0.06	3.3	9.7	16.2
	5	6	53.0	11.90	7.50	-0.56	-0.01	51.2	53.0	63.7
	6	6	83.0	19.90	13.80	-0.71	0.11	80.0	83.0	90.8
	7	7	21.8	8.60	5.20	0.37	0.03	0.0	21.8	26.3
CIM	MODEL			58.23	72.97					
	1	2	0.0	6.49	4.24	-0.32	-0.13	0.0	0.0	7.2
	2	4	83.2	6.44	4.21	-0.32	-0.08	81.2	83.2	87.7
	3	5	54.7	10.29	7.03	0.43	-0.02	32.3	54.7	54.7
	4	6	3.3	19.26	14.63	0.64	-0.15	0.0	3.3	9.3
	5	6	53.0	5.13	3.30	-0.36	-0.01	49.0	53.0	63.7
	6	6	80.6	20.03	15.36	-0.75	0.10	77.6	80.6	84.0
	7	7	21.5	7.36	4.86	0.37	0.04	15.5	21.5	26.3
IM	MODEL			46.17	64.56					
	1	4	87.7	5.63	4.78	-0.34	-0.07	73.8	87.7	87.7
	2	5	54.7	8.80	7.75	0.44	0.06	20.4	54.7	54.7
	3	6	2.2	12.27	11.25	0.55	-0.15	0.0	2.2	13.3
	4	6	83.0	26.24	28.46	-0.90	0.04	77.6	83.0	84.0
	5	7	21.8	4.86	4.09	0.34	-0.01	0.0	21.8	26.3

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

Table 1 continued.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

Table 2. QTL mapping results for parthenocarpic fruit set in the data obtained from experiment 1. The data consists of 201 F_3 families represented by five individuals each. Phenotypic data was collected from nodes 5-30 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			51.03	68.94					
	1	2	0.9	12.66	10.45	-0.76	-0.01	0.0	0.9	4.6
	2	4	83.2	7.33	5.68	-0.51	-0.22	73.8	83.2	87.7
	3	5	32.3	5.56	4.22	0.41	0.38	20.4	32.3	52.6
	4	6	13.3	13.57	11.32	0.85	0.25	3.3	13.3	18.6
	5	6	53.0	10.85	8.77	-0.76	0.11	48.0	53.0	63.7
	6	6	83.0	18.84	16.77	-1.16	0.09	80.6	83.0	90.8
	4×5[u]			3.90	2.90					
CIM	MODEL			42.90	62.58					
	1	2	0.0	9.88	9.51	-0.71	-0.09	0.0	0.0	7.2
	2	4	83.2	7.17	6.68	-0.55	-0.17	81.2	83.2	87.7
	3	5	32.3	4.05	3.64	0.39	0.32	26.4	32.3	52.6
	4	6	2.2	7.07	6.58	0.58	-0.29	0.0	2.2	9.3
	5	6	83.0	29.04	35.37	-1.42	0.02	80.6	83.0	90.8
IM	MODEL			38.24	58.36					
	1	2	0.0	8.60	9.07	-0.70	0.03	0.0	0.0	4.6
	2	4	87.7	6.10	6.25	-0.54	-0.18	58.7	87.7	87.7
	3	6	2.2	6.64	6.84	0.60	-0.23	0.0	2.2	9.3
	4	6	83.0	26.48	34.74	-1.41	-0.02	80.6	83.0	84.0

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

Table 2 continued.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

[u]Epistatic QTL interaction.

Table 3. QTL mapping results for parthenocarpic fruit set in the dataset obtained from the compilation of experiments 2 and 3. The data consists of 205 F_3 families represented by six individuals each. Phenotypic data was collected from nodes 5-30 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			34.69	54.12					
	1	5	54.7	11.40	13.45	0.52	0.15	32.3	54.7	54.7
	2	6	0.0	9.70	11.13	0.49	-0.09	0.0	0.0	13.3
	3	6	80.0	13.10	15.70	-0.60	0.02	76.2	80.0	83.0
	4	7	21.5	9.80	11.25	0.51	0.01	19.1	21.5	33.2
CIM	MODEL			34.69	54.12					
	1	5	54.7	11.40	13.45	0.52	0.15	32.3	54.7	54.7
	2	6	0.0	9.70	11.13	0.49	-0.09	0.0	0.0	9.3
	3	6	80.0	13.10	15.70	-0.60	0.02	75.1	80.0	83.0
	4	7	21.5	9.80	11.25	0.51	0.01	17.2	21.5	26.3
IM	MODEL			34.13	53.55					
	1	5	54.7	10.65	12.56	0.50	0.14	32.3	54.7	54.7
	2	6	2.2	9.16	10.61	0.49	-0.14	0.0	2.2	13.3
	3	6	80.0	12.91	15.63	-0.60	0.02	73.3	80.0	84.0
	4	7	21.8	9.39	10.91	0.50	-0.03	0.0	21.8	26.3

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

Table 3 continued.

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

Table 4. QTL mapping results for parthenocarpic fruit set in the dataset obtained from the compilation of experiments 2 and 3. The data consists of 205 F_3 families represented by six individuals each. Phenotypic data was collected from nodes 5-20 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			46.65	64.93					
	1	4	86.9	5.02	4.19	-0.26	-0.08	84.7	86.9	87.7
	2	5	52.9	12.08	10.93	0.41	0.13	32.3	52.6	54.7
	3	6	0.0	14.75	13.78	0.49	-0.07	0.0	0.0	3.3
	4	6	80.0	23.03	23.76	-0.66	-0.08	77.4	80.0	83.0
	5	7	24.1	9.61	8.45	0.40	0.04	23.0	32.1	42.0
CIM	MODEL			41.76	60.86					
	1	5	52.9	10.77	10.71	0.40	0.18	32.3	52.9	54.7
	2	6	12.6	13.01	13.28	0.48	-0.06	3.3	12.6	16.2
	3	6	80.0	22.12	25.19	-0.68	-0.06	73.3	80.0	80.6
	4	7	24.1	10.36	10.25	0.44	0.02	19.1	24.1	32.1
IM	MODEL			40.77	59.99					
	1	5	52.6	10.78	10.96	0.40	0.17	32.3	52.6	54.7
	2	6	9.7	12.34	12.79	0.47	-0.09	0.0	9.7	16.2
	3	6	80.0	19.92	22.58	-0.64	-0.01	77.6	80.0	84.0
	4	7	21.8	9.52	9.54	0.41	-0.05	0.0	21.8	26.3

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

Table 4 continued.

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

Table 5. QTL mapping results for parthenocarpic fruit set in the dataset obtained from the compilation of experiments 2 and 3. The pooled data consists of 205 F_3 families represented by six individuals each. Phenotypic data was collected from nodes 5-10 of each F_3 plant and used to calculate F_3 family means. Genotypic data was collected from F_2 individuals. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			23.13	40.52					
	1	6	3.3	8.33	12.24	0.26	-0.18	0.0	3.3	9.3
	2	6	80.0	8.65	12.76	-0.29	0.10	63.7	80.0	83.0
	3	7	32.1	11.44	17.43	0.33	0.02	19.1	32.1	42.0
CIM	MODEL			22.17	39.22					
	1	6	3.3	7.23	10.72	0.25	-0.15	0.0	3.3	9.3
	2	6	80.0	7.83	11.69	-0.28	0.11	73.3	80.0	80.6
	3	7	24.1	10.48	16.13	0.33	-0.04	19.1	24.1	32.1
IM	MODEL			21.94	38.92					
	1	6	2.2	7.01	10.42	0.25	-0.15	0.0	2.2	12.6
	2	6	80.0	7.79	11.68	-0.28	0.10	73.3	80.0	90.8
	3	7	24.1	10.21	15.75	0.33	-0.05	17.2	24.1	42.0

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

Table 6. QTL mapping results for seed size. Data was collected from F_3 seed obtained by self-pollination of 205 2A×Gy8 F_2 plants. Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			19.59	35.59					
	1	3	100.2	7.75	12.25	-0.48	0.32	83.2	100.2	106.0
	2	4	10.4	7.53	11.87	-0.40	0.35	0.0	10.4	22.0
	3	5	14.9	6.48	10.09	0.62	-0.12	2.7	14.9	25.7
	4	6	84.0	3.532	5.32	-0.41	-0.07	69.7	84.0	100.5
	1×2[u]			3.22	4.83					
CIM	MODEL			9.42	19.08					
	1	5	14.9	5.60	10.84	0.64	-0.15	2.7	14.9	20.4
	2	6	80.0	3.88	7.38	-0.47	0.05	73.3	80.0	90.8
IM	MODEL			9.42	19.08					
	1	5	14.9	5.60	10.84	0.64	-0.15	2.7	14.9	25.7
	2	6	80.0	3.88	7.38	-0.47	0.05	73.3	80.0	100.5

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

[u]Epistatic QTL interaction.

Table 7. QTL mapping results for seed weight. Data was collected from F_3 seed obtained by self-pollination of 205 2A×Gy8 F_2 plants. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit. All QTL analyses were performed with R/qtl software (Broman et al., 2003).

Method[z]	Optimal QTL Model[y]					QTL Effects[x]		1.5 LOD Interval		
	QTL	CHR	POS[w]	LOD	% Var[v]	Additive	Dominance	Left[w]	QTL[w]	Right[w]
MIM	MODEL			20.43	37.09					
	1	4	6.4	3.16	4.67	-0.07	0.05	0.0	6.4	58.5
	2	5	20.4	6.06	9.27	0.11	0.03	13.9	20.4	25.7
	3	6	80.0	10.56	17.04	-0.09	-0.05	77.6	80.0	84.7
	4	7	23.0	5.79	8.83	-0.04	-0.02	19.1	23.0	26.3
	3×4[u]			5.14	7.78					
CIM	MODEL			11.38	22.75					
	1	5	20.4	6.33	11.92	0.12	0.03	2.7	20.4	25.7
	2	6	80.0	5.93	11.13	-0.11	-0.02	73.3	80.0	90.8
IM	MODEL			13.91	27.06					
	1	4	6.4	3.17	5.43	-0.08	0.03	0.0	6.4	30.2
	2	5	14.9	6.15	10.92	0.13	-0.02	2.7	15.0	25.7
	3	6	80.0	4.72	8.25	-0.10	-0.03	76.2	80.0	100.5

[z]QTL mapping method: MIM = Multiple Interval Mapping; CIM = Composite Interval Mapping; IM = Interval Mapping.

[y]Best fit model selected to fit all significant QTL and QTL interactions.

[x]QTL effects calculated as: Additive (a) = (μAA - μBB)/2 and Dominance (d) = μAB - (μAA + μBB)/2.

[w]Positions correspond with the linkage map presented in Chapter 2. Map distances were calculated using the Kosambi function in JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

[v]Percent of the phenotypic variation (heritability due to the QTL) explained by the full model and each individual QTL. Follows the equation: $h^2 = (2a^2 + d^2)/(2a^2 + d^2 + 4\sigma^2)$ where σ^2 represents the residual variation.

[u]Epistatic QTL interaction.

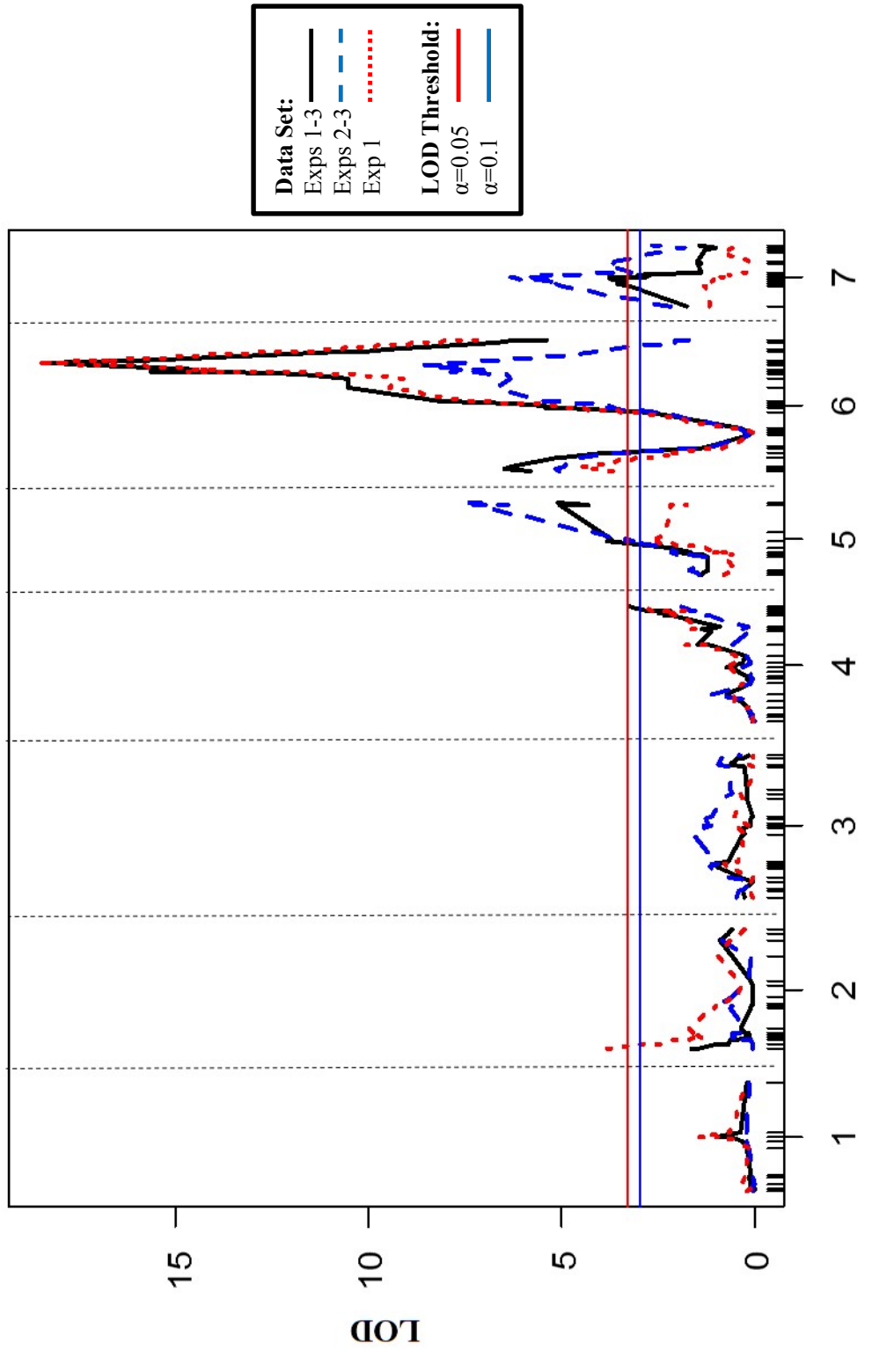


Figure 1. Plot of genome wide LOD curves obtained by interval mapping from data collected for experiments 1-3. After experiment 1 was observed to differ from experiments 2 and 3 in Chapter 1, a separate QTL analysis was conducted. Pooled data from experiments 1-3 is compared with data from experiment 1 and a dataset composed of experiments 2 and 3 combined. QTL analyses were performed with R/qtl software (Broman et al., 2003).

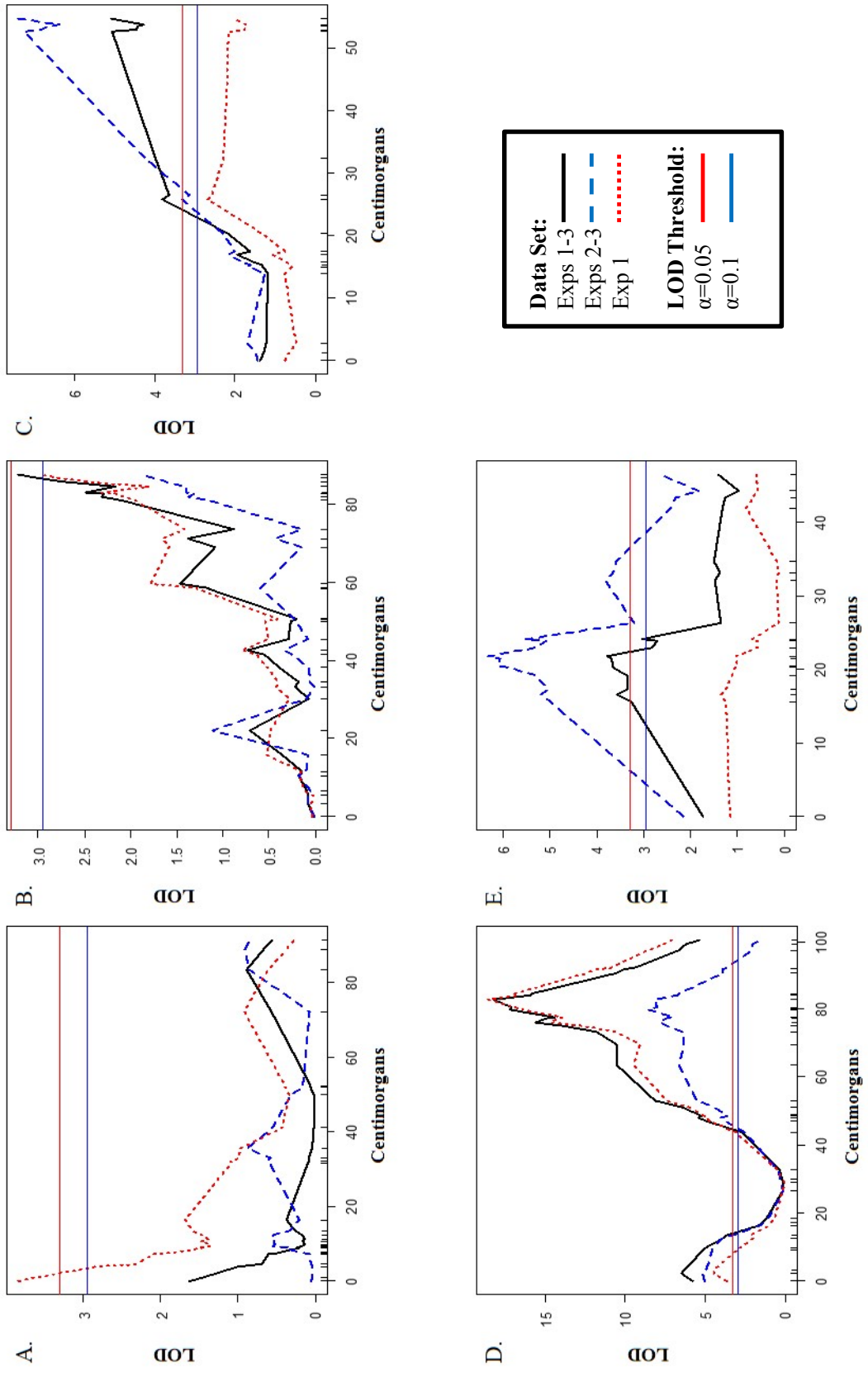


Figure 2. Plot of LOD curves for chromosomes 2(A), 4(B), 5(C), 6(D), and 7(E) obtained by interval mapping from data collected for experiments 1-3. After experiment 1 was observed to differ from experiments 2 and 3 in Chapter 1, a separate QTL analysis was conducted. Pooled data from experiments 1-3 is compared with data from experiment 1 and a dataset composed of experiments 2 and 3 combined. QTL analyses were performed with R/qtl software (Broman et al., 2003).

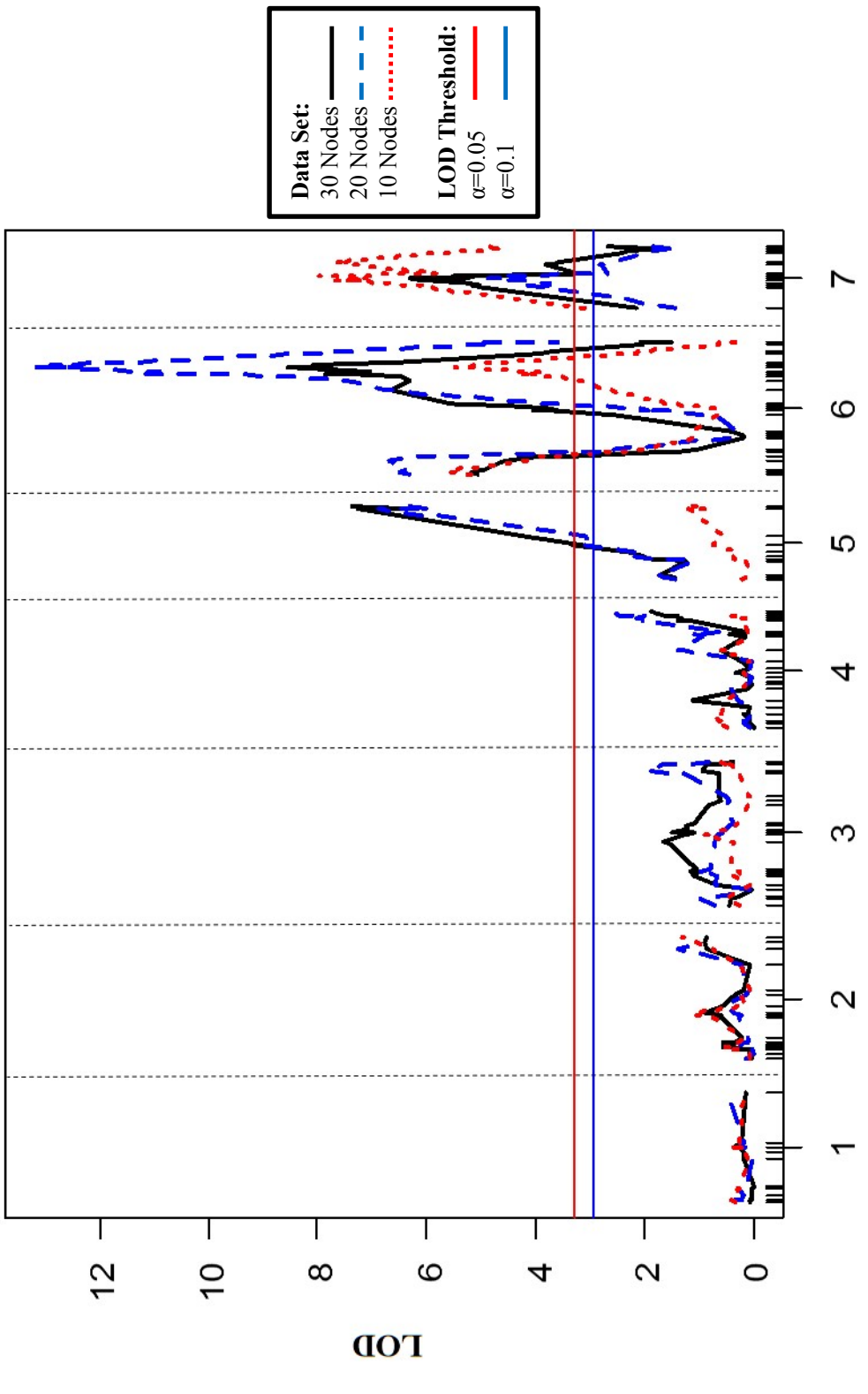


Figure 3. Plot of genome wide LOD curves obtained by interval mapping from data collected from the first 10, 20, and 30 nodes of each F_3 plant. For this comparison only data from experiments 2 and 3 were included due to the delayed fruit set observed in experiment 1. QTL analyses were performed with R/qtl software (Broman et al., 2003).

Figure 4. Plot of LOD curves for chromosomes 5(A), 6(B), and 7(C) obtained by interval mapping from data collected from the first 10, 20, and 30 nodes of each F_3 plant. For this comparison only data from experiments 2 and 3 were included due to the delayed fruit set observed in experiment 1. QTL analyses were performed with R/qtl software (Broman et al., 2003).

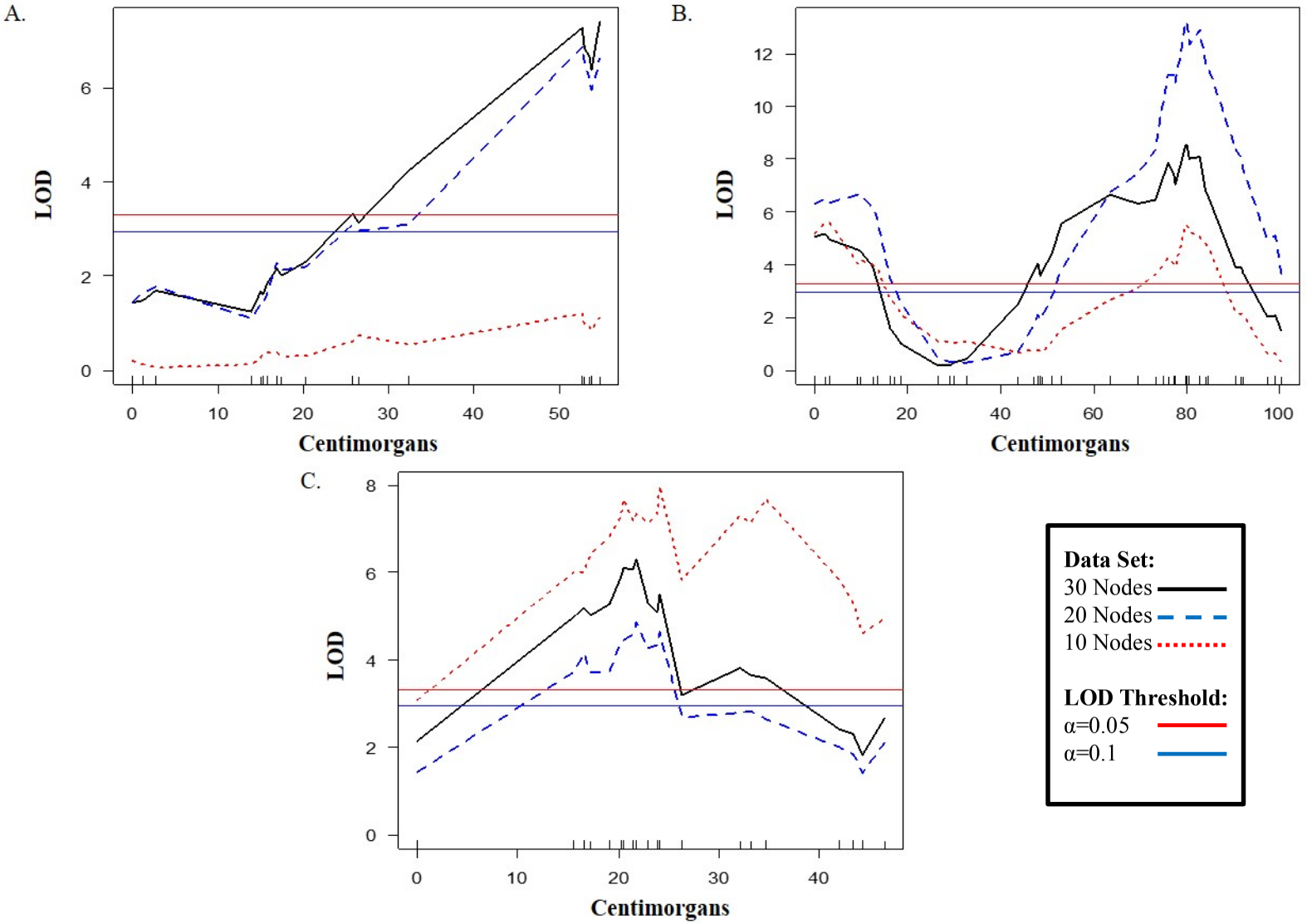


Figure 5. Plot of genome wide LOD curves obtained by interval mapping from data collected for the seed size of each F_2 plant. Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds. QTL analyses were performed with R/qtl software (Broman et al., 2003).

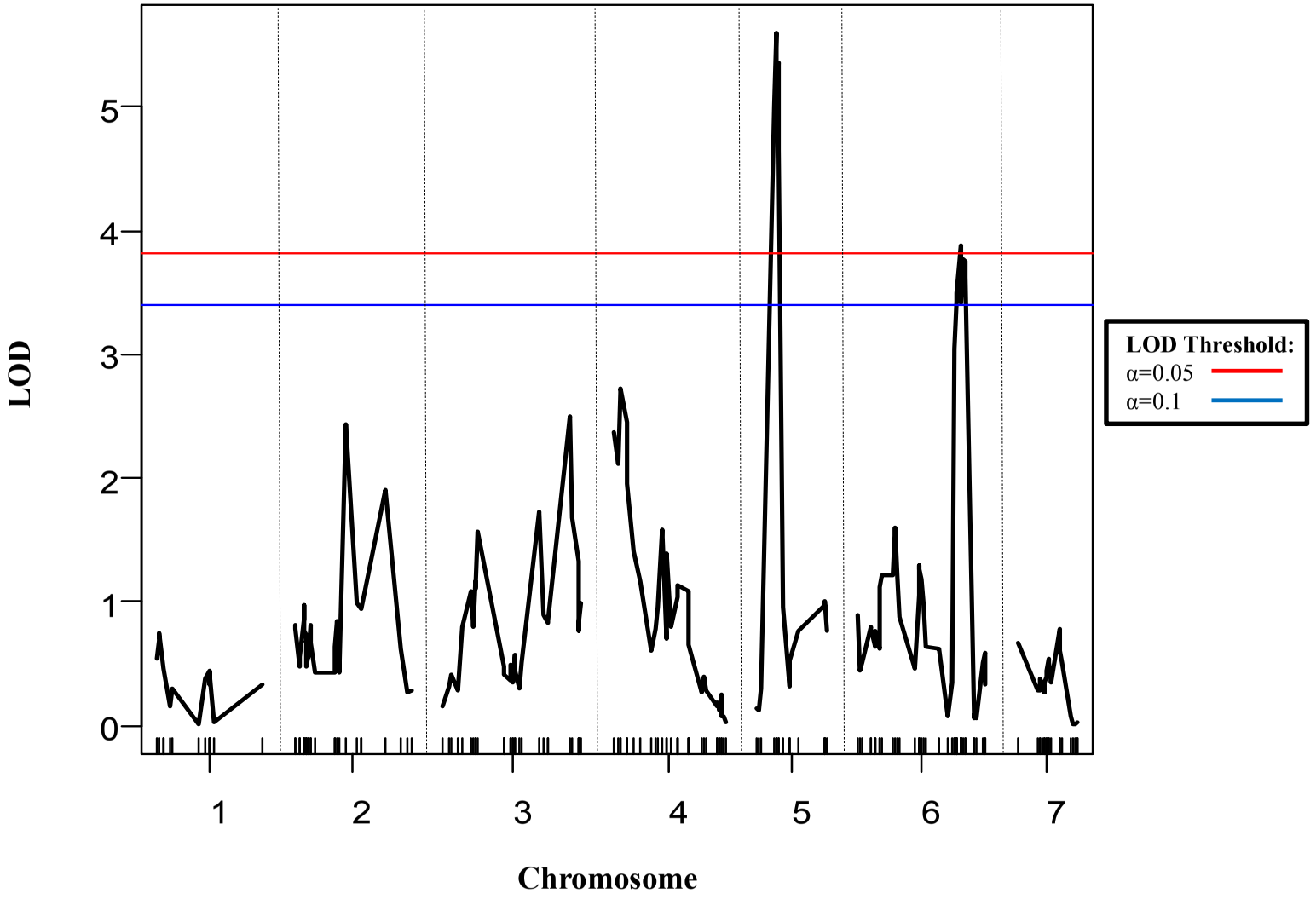


Figure 6. Plot of LOD curves for chromosomes 5(A) and 6(B) obtained by interval mapping from data collected for the seed size and weight of each F_2 plant. Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit. QTL analyses were performed with R/qtl software (Broman et al., 2003).

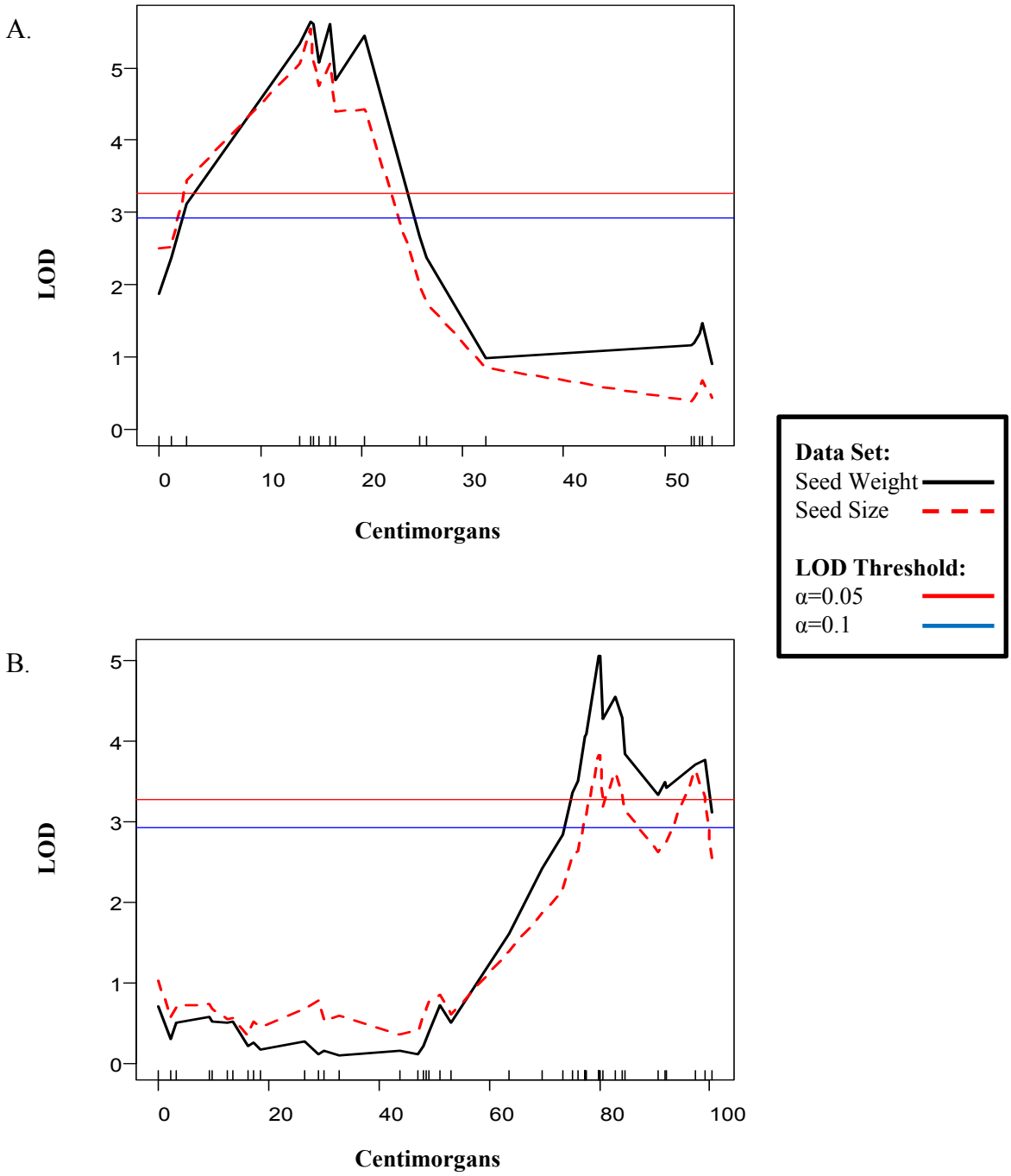


Figure 7. Plot of genome wide LOD curves obtained by interval mapping from data collected for the seed weight of each F_2 plant. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit. QTL analyses were performed with R/qtl software (Broman et al., 2003).

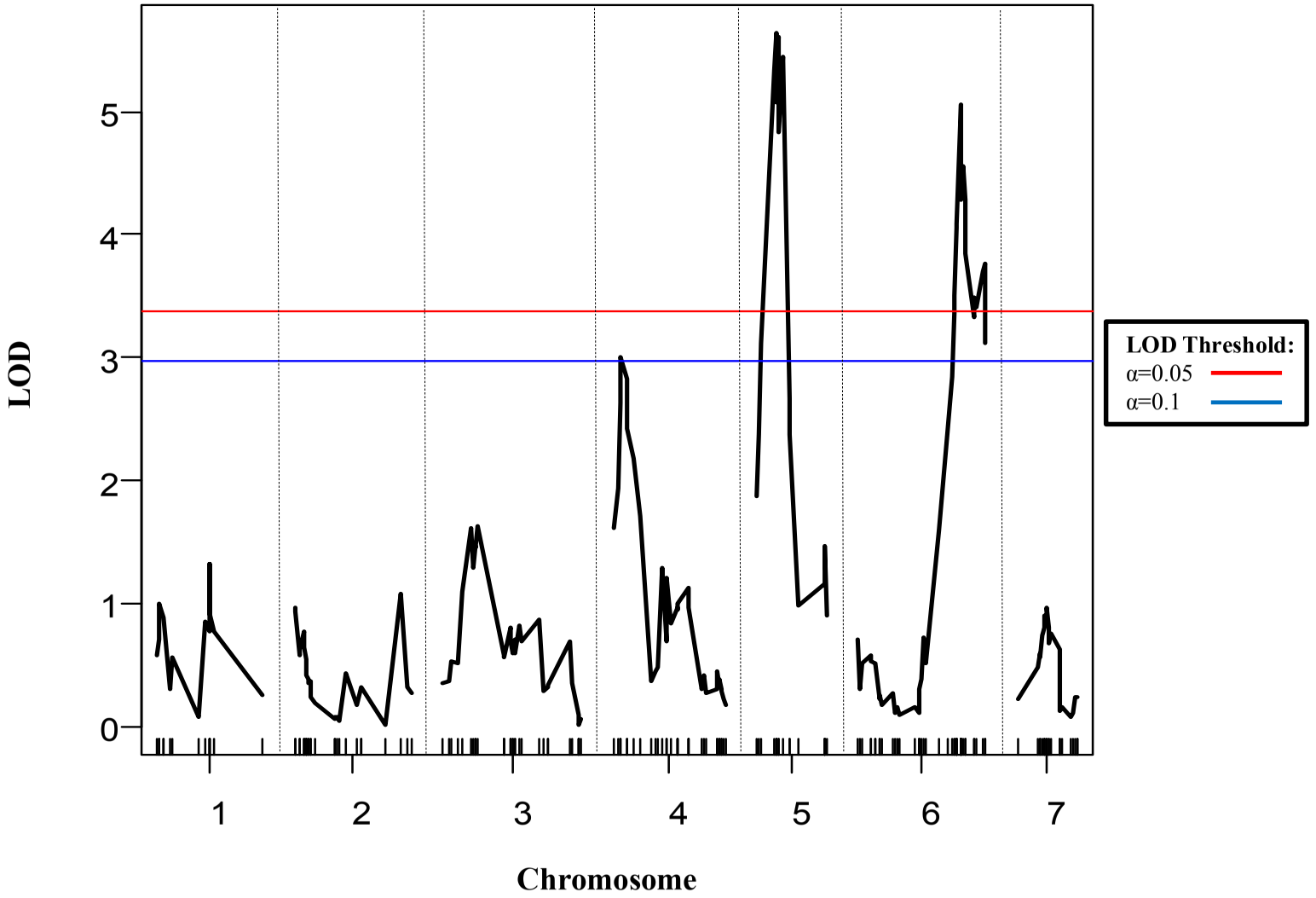
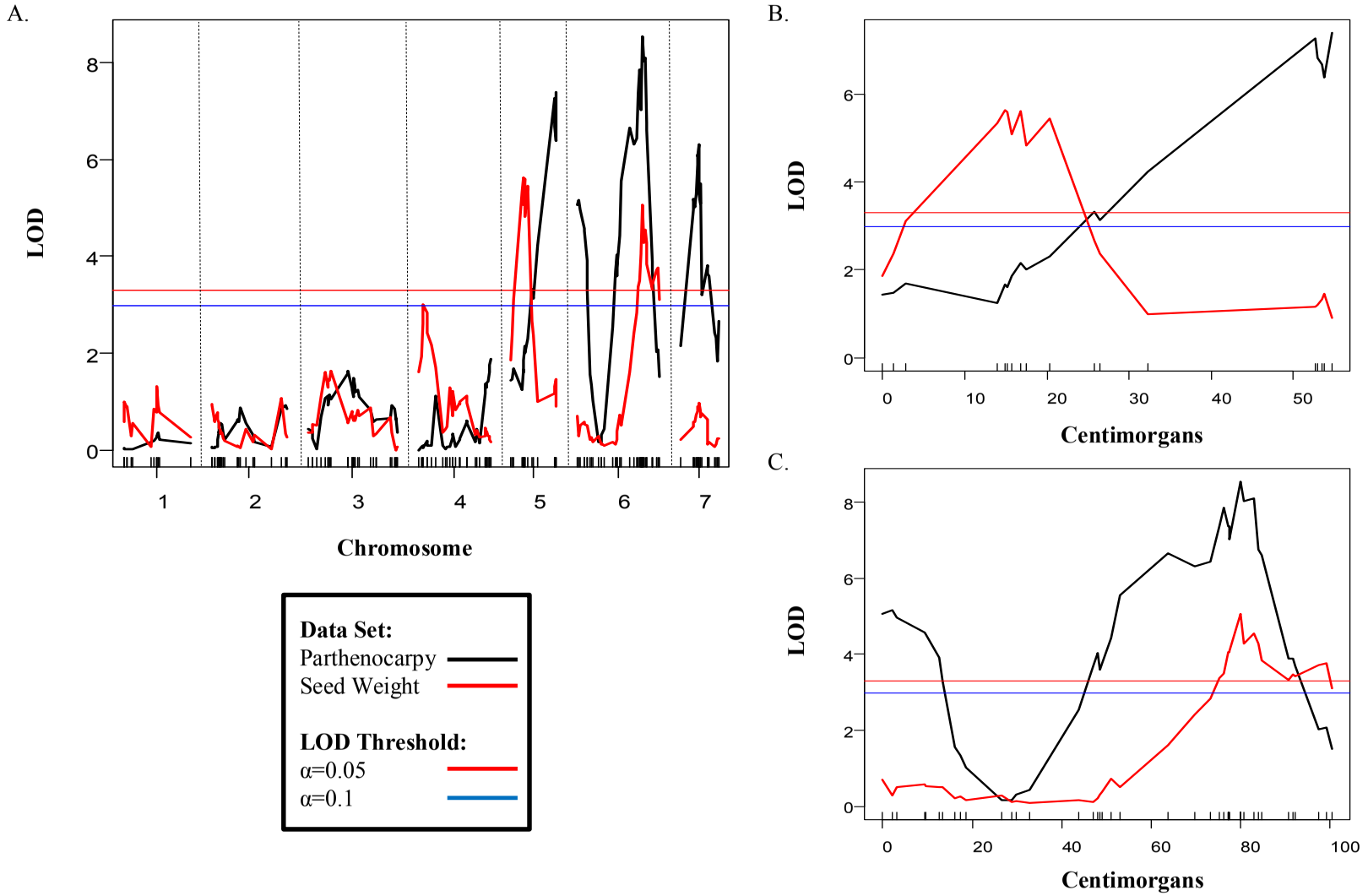
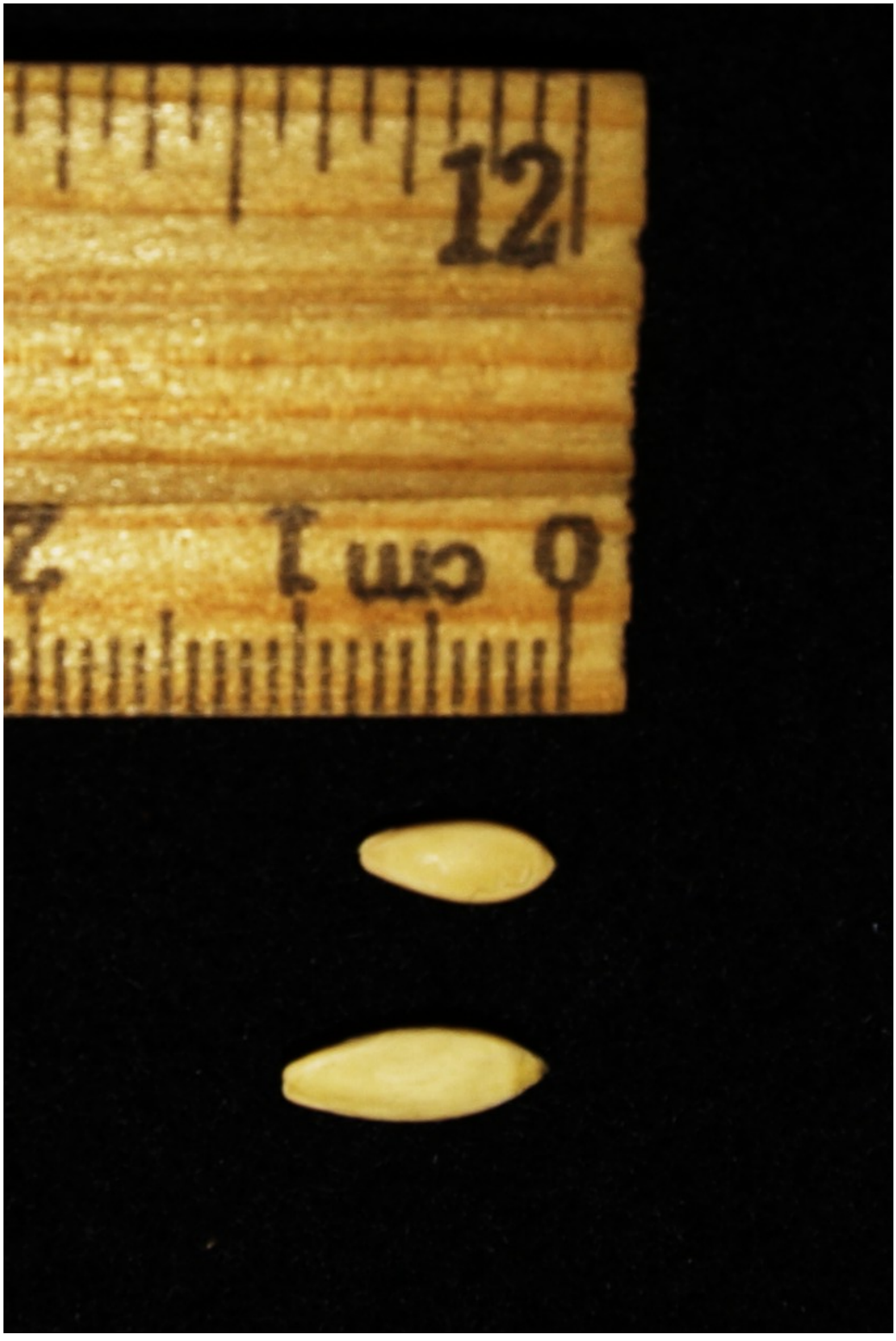


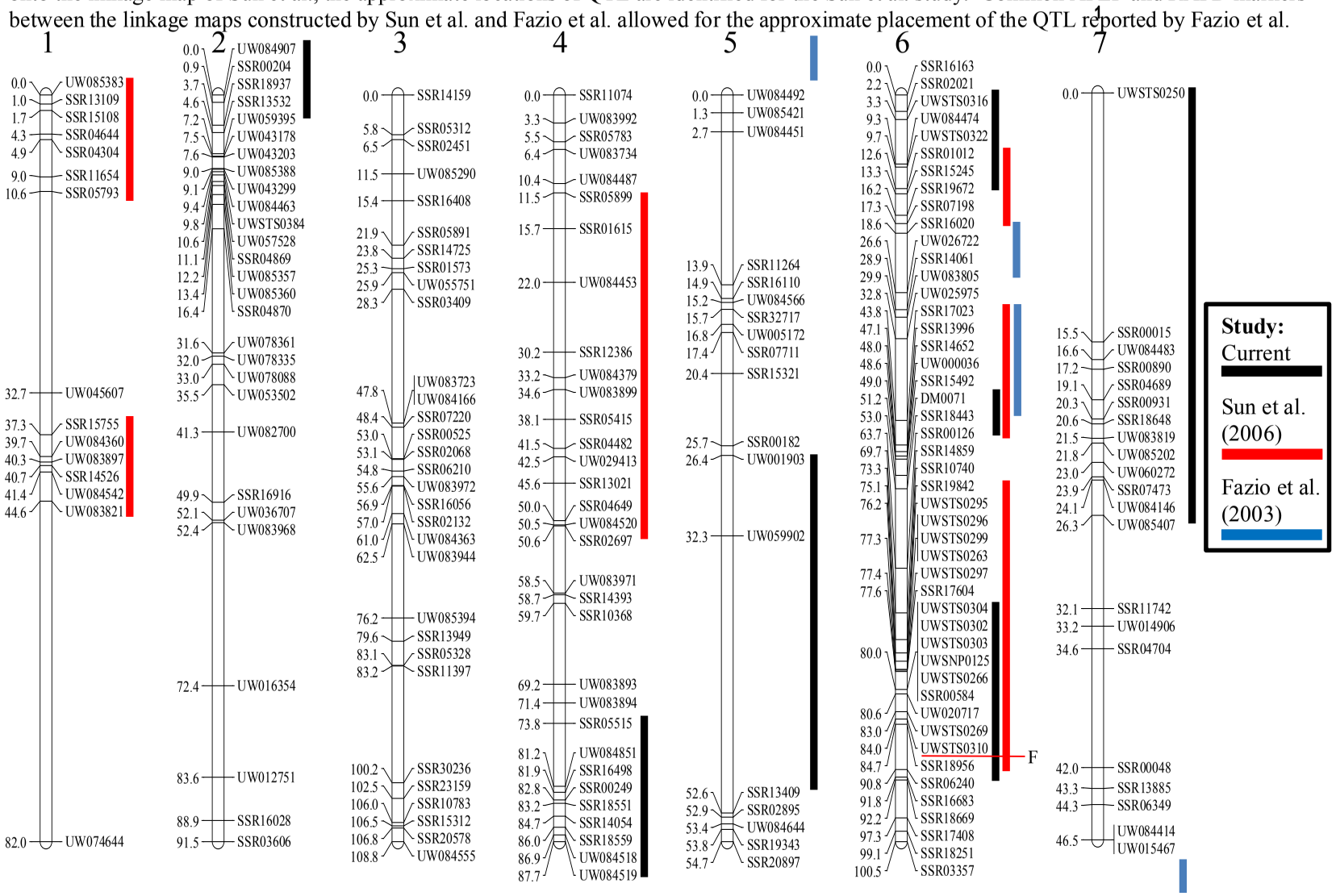
Figure 8. Plot of LOD curves for comparison of parthenocarpic fruit set and seed weight traits. Data for parthenocarpic fruit set is obtained from a data set consisting of data from experiments 2 and 3 combined. Parthenocarpic fruit set is measured as the number of parthenocarpic fruits initiating growth on the first 30 nodes of each F_3 plant. Data is presented as F_3 family means. Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit of each F_2 plant. A: LOD curves of whole genome scans of both traits obtained by interval mapping. B and C: LOD curves obtained by interval mapping for both traits on individual chromosomes 5(B) and 6(C).



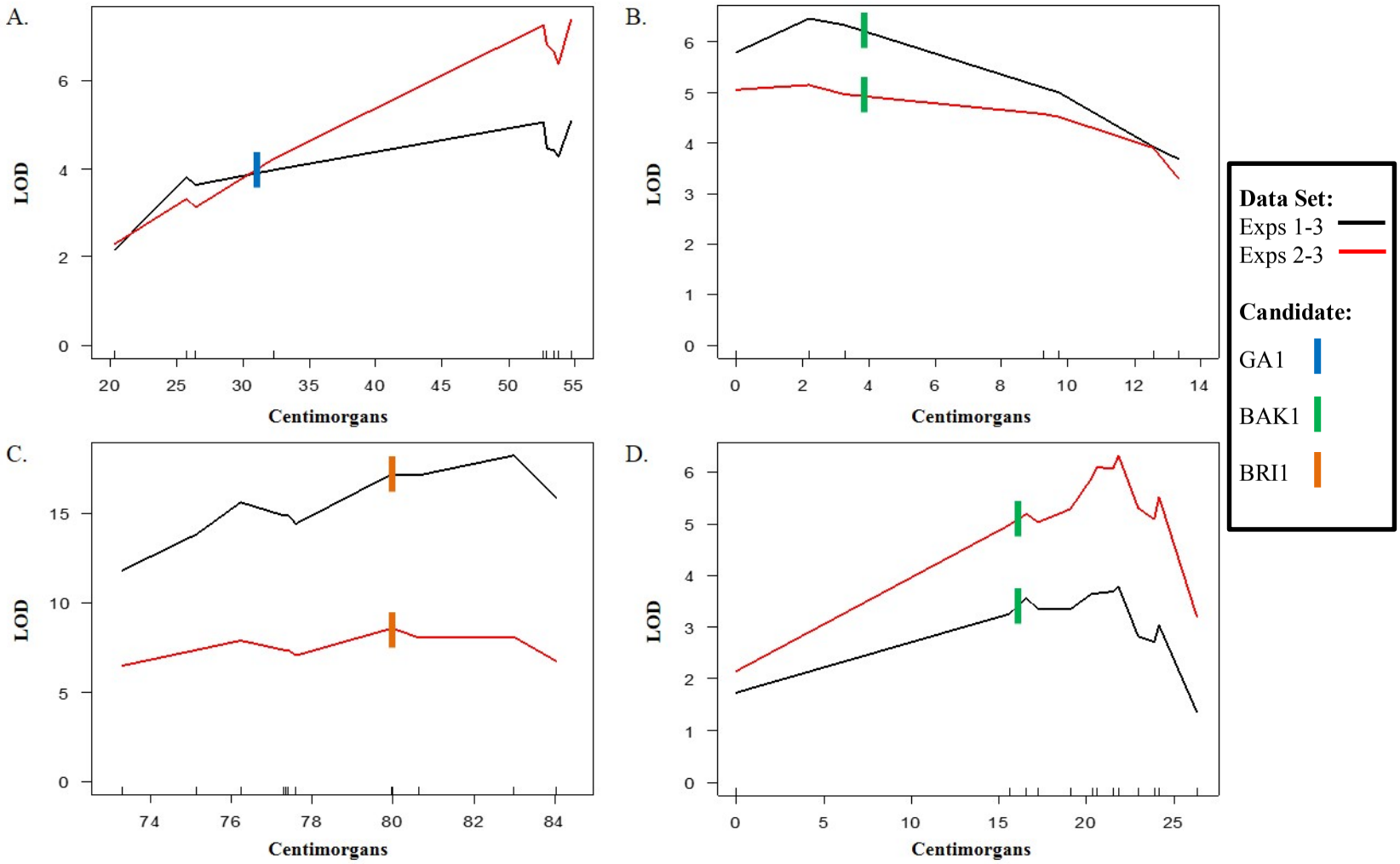


Addendum 1. Photograph depicting the difference in seed size between the parental inbred lines‘2A’ (right) and ‘Gy8’ (left). The seeds shown in this photo are healthy and fully developed seeds and are representative of seeds from each parental inbred line.

Addendum 2. The locations of QTL identified by the current study for "parthenocarpic fruit set" are compared with the locations of QTL identified for "parthenocarpic yield" by Sun et al. (2006) and "number of fruits per plant" by Fazio et al. (2003). All QTL are placed onto the linkage map constructed for the 2A×Gy8 $F_{2:3}$ mapping population used in the current study. With the use of the common SSR markers placed onto the linkage map of Sun et al., the approximate locations of QTL are identified for the Sun et al. study. Common AFLP and RAPD markers between the linkage maps constructed by Sun et al. and Fazio et al. allowed for the approximate placement of the QTL reported by Fazio et al.



Addendum 3. Close up view of the LOD curves within the 1.5 LOD confidence interval for each of the four consensus QTL associated with parthenocarpic fruit set on chromosomes 5 (A), 6 (B and C), and 7 (D). Colored dashes mark the approximate location of each candidate gene. Candidate gene locations were derived by a BLASTn search of the predicted candidate gene sequence to the Gy14 Draft Genome Assembly Version 1.0 (Yang et al., 2012). These positions were then compared to known molecular marker positions within the assembly. Molecular markers are marked as tick marks on the x-axis of each graph (refer to Chapter 2 Table 5 for marker names and physical position within the assembly).



Addendum 4. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth6.2* QTL (sequence between molecular markers SSR17604 and UWSTS0310) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
protein brassinosteroid insensitive 1 (BRI1) mRNA	**1299**	**1299**	**70%**	**NM_120100.2**
pentatricopeptide repeat-containing protein mRNA	922	922	68%	NM_117439.2
pyridoxal phosphate (PLP)-dependent transferases superfamily protein mRNA	917	1178	72%	NM_119873.1
catalytic/ pyridoxal phosphate binding protein mRNA	877	1104	72%	NM_127916.2
catalytic/ pyridoxal phosphate binding protein mRNA	805	1137	73%	NM_126094.3
actin 7 mRNA	695	1341	85%	NM_121018.3
actin 8 mRNA	650	1207	83%	NM_103814.3
actin 2 mRNA	648	1179	79%	NM_180280.1
plant glycogenin-like starch initiation protein 1 mRNA	643	767	75%	NM_001035645.2
plant glycogenin-like starch initiation protein 1 mRNA	643	767	75%	NM_112752.3
protein kinase-like protein ABC1K10 mRNA	630	1013	74%	NM_101012.4
actin 2 mRNA	628	1226	83%	NM_112764.3
putative UDP-glucuronate:xylan alpha-glucuronosyltransferase 3 mRNA	618	785	75%	NM_106363.4
uncharacterized protein mRNA	600	664	70%	NM_114872.2
actin 3 mRNA	594	1157	81%	NM_115235.3
actin-11 mRNA	580	1177	81%	NM_112046.3
PS II oxygen-evolving complex 1 mRNA	578	658	77%	NM_126055.3
actin 1 mRNA	576	1148	81%	NM_179953.2
actin 1 mRNA	576	1148	81%	NM_001036427.2
oxygen-evolving enhancer protein 1-2 mRNA	547	547	76%	NM_114942.2
E3 ubiquitin-protein ligase RKP mRNA	536	1336	71%	NM_179689.2
wall-associated receptor kinase-like 14 mRNA	522	594	73%	NM_179710.2

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
wall-associated receptor kinase-like 14 mRNA	522	594	73%	NM_127909.1
actin-12 mRNA	495	1020	78%	NM_114519.2
actin 4 mRNA	493	995	78%	NM_125328.3
actin 4 mRNA	493	995	78%	NM_001085300.1
uncharacterized protein mRNA	486	486	69%	NM_114894.2
putative galacturonosyltransferase-like 1 mRNA	479	479	72%	NM_101787.2
uncharacterized protein mRNA	468	468	69%	NM_114877.3
uncharacterized protein mRNA	457	457	69%	NM_114873.2
receptor-like protein kinase BRI1-like 3 mRNA	444	494	73%	NM_112183.2
Actin-like ATPase superfamily protein mRNA	441	809	76%	NM_129773.1
geranylgeranyl pyrophosphate synthase 1 mRNA	428	428	73%	NM_119845.3
uncharacterized protein mRNA	425	425	68%	NM_126102.2
geranylgeranyl pyrophosphate synthase 7 mRNA	414	414	72%	NM_127418.1
uncharacterized protein mRNA	408	408	74%	NM_126121.2
ABC1 domain-containing kinase mRNA	407	572	72%	NM_202336.2
ABC1 domain-containing kinase mRNA	407	563	72%	NM_104846.3
serine/threonine-protein kinase BRI1-like 1 mRNA	401	448	72%	NM_001124029.1
serine/threonine-protein kinase BRI1-like 1 mRNA	401	448	72%	NM_104437.2
protein HOS3-1 mRNA	401	401	71%	NM_119847.3
putative galacturonosyltransferase-like 2 mRNA	381	381	70%	NM_114936.4
general control non-repressible 5 mRNA	360	1238	74%	NM_125882.2
uncharacterized protein mRNA	356	644	69%	NM_114938.3
actin family protein mRNA	354	740	73%	NM_180032.1
serine/threonine-protein kinase BRI1-like 2 mRNA	349	392	71%	NM_126256.3
Autophagy-related protein 13 mRNA	327	454	68%	NM_112763.4
protein IRREGULAR XYLEM 15 mRNA	327	388	72%	NM_114882.1
uncharacterized protein mRNA	320	320	74%	NM_112766.1
ABC transporter F family member 2 mRNA	316	1139	80%	NM_121030.2

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
actin 9 mRNA	313	588	71%	NM_129772.1
extra-large G-protein 1 mRNA	311	867	77%	NM_127910.2
protein BRUSHY 1 mRNA	307	775	69%	NM_112759.4
geranylgeranyl pyrophosphate synthase 4 mRNA	306	306	70%	NM_127420.1
uncharacterized protein mRNA	295	295	77%	NM_203265.2
wall-associated receptor kinase-like 21 mRNA	295	295	70%	NM_126077.4
uncharacterized protein mRNA	295	295	77%	NM_125879.3
GNS1/SUR4 membrane-like protein mRNA	291	291	70%	NM_106157.1
geranylgeranyl pyrophosphate synthase 11 mRNA	289	289	69%	NM_113869.1
putative geranylgeranyl pyrophosphate synthase 8 mRNA	280	280	68%	NM_112311.1
geranylgeranyl pyrophosphate synthase 3 mRNA	279	279	68%	NM_112315.3
ribonucleoside-diphosphate reductase large subunit mRNA	266	1611	82%	NM_127748.3
FAD-dependent oxidoreductase-like protein mRNA	266	442	68%	NM_126129.3
O-Glycosyl hydrolases family 17 protein mRNA	260	365	68%	NM_125875.2
geranylgeranyl pyrophosphate synthase 9 mRNA	259	259	67%	NM_112313.2
geranylgeranyl pyrophosphate synthase 2 mRNA	255	255	67%	NM_127943.2
BEL1-like homeodomain 4 mRNA	246	578	74%	NM_179713.2
BEL1-like homeodomain 4 mRNA	246	578	74%	NM_127939.2
BEL1-like homeodomain 4 mRNA	246	578	74%	NM_001036327.1
CDPK-related kinase mRNA	242	1097	78%	NM_114913.3
uncharacterized protein mRNA	241	475	73%	NM_001084797.2
phototropic-responsive NPH3 family protein mRNA	239	540	65%	NM_126054.2
geranylgeranyl pyrophosphate synthase 12 mRNA	235	235	67%	NM_114027.2
RING/U-box superfamily protein mRNA	232	232	71%	NM_127941.3
probable beta-1,4-xylosyltransferase IRX14H mRNA	219	429	68%	NM_126123.4
putative galacturonosyltransferase-like 3 mRNA	219	219	67%	NM_101196.2
probable beta-1,4-xylosyltransferase IRX14 mRNA	214	502	68%	NM_119853.3
uncharacterized protein mRNA	214	459	69%	NM_114878.2

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
protein SKU5 similar 17 mRNA	214	1016	77%	NM_126091.3
transcription factor HEC3 mRNA	214	214	76%	NM_121012.1
geranylgeranyl pyrophosphate synthase 10 mRNA	212	212	67%	NM_112907.1
CDPK-related kinase 2 mRNA	210	485	76%	NM_112797.4
cytochrome P450, family 96, subfamily A, polypeptide 1 mRNA	205	1747	68%	NM_127882.2
uridylate kinase-like protein mRNA	203	594	79%	NM_112754.3
serine/threonine-protein kinase/endoribonuclease IRE1a mRNA	199	700	76%	NM_127306.3
extra-large GTP-binding protein 2 mRNA	196	196	71%	NM_119604.3
CDPK-related kinase 7 mRNA	194	468	76%	NM_115535.2
protein PHOSPHATASE AND TENSIN HOMOLOG mRNA	194	622	84%	NM_114871.2
putative galacturonosyltransferase-like 7 mRNA	187	235	70%	NM_116131.4
nucleotide sugar transporter-KT 1 mRNA	187	558	76%	NM_179196.1
nucleotide sugar transporter-KT 1 mRNA	187	558	76%	NM_120099.3
RING/U-box superfamily protein mRNA	185	185	69%	NM_101788.3
ethylene-responsive transcription factor RAP2-10 mRNA	185	240	74%	NM_119854.2
ethylene-responsive transcription factor ERF010 mRNA	185	251	74%	NM_126119.1
uncharacterized protein mRNA	185	314	67%	NM_114876.2
CDPK-related kinase 3 mRNA	183	435	74%	NM_130235.4
P-loop containing nucleoside triphosphate hydrolases superfamily mRNA	181	346	75%	NM_114922.4
putative galacturonosyltransferase-like 10 mRNA	176	226	70%	NM_113753.4
ubiquinol-cytochrome C chaperone family protein mRNA	174	174	82%	NM_124501.3
nucleotide-sugar transporter family protein mRNA	174	484	75%	NM_103124.3

[z]Description of annotated gene function provided by NCBI.

[y]Calculated from the sum of the match rewards and mismatch/open gap penalties for each segment.

[x]Sum of alignment scores of all segments.

[w]Highest percent identity for a set of aligned segments.

Addendum 5. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth6.1* QTL (sequence between molecular markers UWSTS0316 and SSR19672) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
Leu-rich receptor Serine/threonine protein kinase BAK1 mRNA	**408**	**1139**	**83%**	**NM_119497.4**
Protein phosphatase 2A B' alpha mRNA	**398**	**462**	**73%**	**NM_120427.2**
cell division cycle protein 48-related protein mRNA	1056	1725	74%	NM_100472.1
Clathrin, heavy chain mRNA	1054	4669	83%	NM_111950.2
Clathrin, heavy chain mRNA	1025	4687	82%	NM_111688.6
putative ATP-dependent RNA helicase mRNA	1009	1009	73%	NM_123347.1
bifunctional alpha-l-arabinofuranosidase/beta-d-xylosidase mRNA	924	1237	73%	NM_105527.4
adenine/guanine permease AZG1 mRNA	906	906	73%	NM_111933.2
cytochrome P450, family 86, subfamily A, polypeptide 4 mRNA	881	881	74%	NM_100042.3
cytochrome P450 86A2 mRNA	834	834	73%	NM_116260.3
phospholipid-transporting ATPase 1 mRNA	787	1618	73%	NM_120575.2
cytochrome P450, family 86, subfamily A, polypeptide 8 mRNA	782	782	72%	NM_130160.2
C2 calcium/lipid-binding plant phosphoribosyltransferase family protein mRNA	760	888	68%	NM_117230.3
leucine-rich repeat protein kinase-like protein mRNA	719	1179	72%	NM_115495.3
protein phosphatase 2A regulatory subunit B' eta mRNA	675	864	76%	NM_001035693.2
protein phosphatase 2A regulatory subunit B' eta mRNA	675	864	76%	NM_113506.1
protein phosphatase 2A B'theta mRNA	659	659	75%	NM_101216.1
protein phosphatase 2A B'theta mRNA	659	659	75%	NM_202087.1
C2 domain-containing plant phosphoribosyltransferase-like protein mRNA	657	657	67%	NM_115650.4
cytochrome P450, family 86, subfamily A, polypeptide 7 mRNA	639	639	71%	NM_105048.2
putative inorganic phosphate transporter 1-5 mRNA	625	625	70%	NM_128843.3
protein ethylene insensitive 3 mRNA	623	623	77%	NM_112968.3
ethylene insensitive 3-like 1 protein mRNA	612	668	77%	NM_128263.4

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
inorganic phosphate transporter 1-4 mRNA	572	572	69%	NM_129452.3
cysteine-rich peptide family protein mRNA	571	571	72%	NM_129528.2
bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase mRNA	567	986	74%	NM_100043.4
protein phosphatase 2A B'gamma mRNA	554	554	73%	NM_117630.2
protein phosphatase 2A B'gamma mRNA	554	554	73%	NM_179059.1
phosphate transporter 1;7 mRNA	540	540	68%	NM_115327.3
putative inorganic phosphate transporter 1-3 mRNA	538	538	69%	NM_123702.1
ribonuclease II/R family protein mRNA	526	574	68%	NM_106417.2
trigalactosyldiacylglycerol 1 mRNA	522	522	75%	NM_202139.2
trigalactosyldiacylglycerol 1 mRNA	522	522	75%	NM_101836.3
trigalactosyldiacylglycerol 1 mRNA	522	522	75%	NM_202140.1
POZ/BTB containin G-protein 1 mRNA	520	939	75%	NM_116025.3
POZ/BTB containin G-protein 1 mRNA	520	939	75%	NM_180402.1
laccase 5 mRNA	508	760	71%	NM_129597.3
inorganic phosphate transporter 1-1 mRNA	508	508	68%	NM_123701.3
ABC transporter C family member 2 mRNA	506	2136	75%	NM_129020.3
GDP-mannose pyrophosphorylase/ mannose-1-pyrophosphatase mRNA	499	901	76%	NM_129535.3
PLAC8 family protein mRNA	497	497	70%	NM_120617.2
protein phosphatase 2A B'zeta mRNA	497	497	72%	NM_113060.4
cytochrome P450, family 77, subfamily A, polypeptide 4 mRNA	489	489	68%	NM_120548.2
probable mannose-1-phosphate guanylyltransferase 2 mRNA	488	822	76%	NM_115416.2
laccase 12 mRNA	480	917	72%	NM_120621.1
Fatty acid/sphingolipid desaturase mRNA	475	475	68%	NM_116023.2
serine/threonine-protein phosphatase PP1 isozyme 4 mRNA	471	730	78%	NM_129543.2
60S ribosomal protein L3-1 mRNA	457	1153	85%	NM_103469.3
60S ribosomal protein L3-1 mRNA	457	939	85%	NM_001084202.1
60S ribosomal protein L3-1 mRNA	457	1153	85%	NM_001036069.1
60S ribosomal protein L3-1 mRNA	457	1153	85%	NM_202237.1

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
protein LIGHT-RESPONSE BTB 1 mRNA	448	779	73%	NM_130189.2
bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase mRNA	446	865	72%	NM_116264.5
calcium-dependent lipid-binding phosphoribosyltransferase-like protein mRNA	439	439	65%	NM_121300.2
glutathione S-conjugate transporting ATPase mRNA	437	1844	74%	NM_001036039.1
glutathione S-conjugate transporting ATPase mRNA	437	1844	74%	NM_102777.2
HXXXD-type acyl-transferase-like protein mRNA	434	434	67%	NM_129556.3
pentatricopeptide repeat-containing protein mRNA	428	428	68%	NM_101211.3
ribulose bisphosphate carboxylase/oxygenase activase mRNA	425	944	80%	NM_179989.2
serine/threonine protein phosphatase 2A B'55 delta mRNA	425	512	70%	NM_113507.2
ribulose bisphosphate carboxylase/oxygenase activase mRNA	425	944	80%	NM_179990.1
ribulose bisphosphate carboxylase/oxygenase activase mRNA	425	944	80%	NM_129531.2
serine/threonine-protein phosphatase PP1 isozyme 2 mRNA	421	635	77%	NM_180887.4
serine/threonine-protein phosphatase PP1 isozyme 2 mRNA	421	635	77%	NM_125306.2
serine/threonine-protein phosphatase PP1 isozyme 2 mRNA	421	635	77%	NM_001037026.1
putative alpha-xylosidase 2 mRNA	417	938	74%	NM_114463.1
cytochrome P450, family 77, subfamily A, polypeptide 9 mRNA	410	410	67%	NM_120545.1
Splicing factor U2af large subunit B mRNA	407	756	80%	NM_104771.3
cellulose synthase A catalytic subunit 5 [UDP-forming] mRNA	405	2083	77%	NM_121024.2
serine/threonine-protein phosphatase PP1 isozyme 1 mRNA	401	550	76%	NM_128494.4
serine/threonine-protein phosphatase PP1 isozyme 5 mRNA	399	600	76%	NM_114549.3
uncharacterized protein mRNA	399	1066	77%	NM_115472.4
cellulose synthase A catalytic subunit 2 [UDP-forming] mRNA	398	2211	75%	NM_120095.3
histone H3.1 mRNA	392	392	81%	NM_125934.2
uncharacterized protein mRNA	390	2337	75%	NM_116066.3
Splicing factor U2af large subunit A mRNA	387	899	79%	NM_202966.3
Splicing factor U2af large subunit A mRNA	387	906	79%	NM_179178.2
Splicing factor U2af large subunit A mRNA	387	902	79%	NM_119833.3
cytochrome P450, family 77, subfamily A, polypeptide 6 mRNA	387	387	66%	NM_111893.3

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
40S ribosomal protein S17-3 mRNA	379	379	83%	NM_111897.3
40S ribosomal protein S17-2 mRNA	378	378	82%	NM_126548.4
40S ribosomal protein S17-2 mRNA	378	378	82%	NM_001036248.1
cellulose synthase A catalytic subunit 6 [UDP-forming] mRNA	374	2062	76%	NM_125870.2
40S ribosomal protein S17-1 mRNA	372	372	84%	NM_126472.3
putative nitrite transporter mRNA	367	644	70%	NM_105528.4
lysine histidine transporter 5 mRNA	367	702	75%	NM_105432.1
histone H3 mRNA	365	365	80%	NM_121078.2
pollen-specific leucine-rich repeat extensin-like protein 1 mRNA	363	1850	76%	NM_112788.2
40S ribosomal protein S17-4 mRNA	360	360	81%	NM_180434.2
alpha/beta-Hydrolases superfamily protein mRNA	360	505	72%	NM_202761.2
40S ribosomal protein S17-4 mRNA	360	360	81%	NM_120562.3
Fatty acid/sphingolipid desaturase mRNA	360	458	69%	NM_130183.3
40S ribosomal protein S17-4 mRNA	360	360	81%	NM_001036758.1
40S ribosomal protein S17-4 mRNA	360	360	81%	NM_001036757.1
alpha/beta-Hydrolases superfamily protein mRNA	360	505	72%	NM_116274.3
histone H3 mRNA	358	358	79%	NM_113651.2
uncharacterized protein mRNA	354	1025	75%	NM_129600.3
uncharacterized protein mRNA	354	1025	75%	NM_179998.1
40S ribosomal protein S15a-1 mRNA	351	351	80%	NM_202054.3
40S ribosomal protein S15a-1 mRNA	351	351	80%	NM_100651.4

[z]Description of annotated gene function provided by NCBI.

[y]Calculated from the sum of the match rewards and mismatch/open gap penalties for each segment.

[x]Sum of alignment scores of all segments.

[w]Highest percent identity for a set of aligned segments.

Addendum 6. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth7.1* QTL (sequence between molecular markers SSR00015 and UW085407) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
Leu-rich receptor Serine/threonine protein kinase BAK1 mRNA	**318**	**1016**	**78%**	**NM_119497.4**
ribosomal protein S5/Elongation factor G/III/V family protein mRNA	1301	1596	75%	NM_113198.2
pentatricopeptide repeat-containing protein mRNA	1159	1223	73%	NM_120364.1
carbamoyl phosphate synthetase B mRNA	1027	2488	75%	NM_102730.1
probable galacturonosyltransferase 10 mRNA	976	976	76%	NM_127647.2
ABC transporter C family member 3 mRNA	825	1470	73%	NM_202570.1
ABC transporter C family member 3 mRNA	825	1828	73%	NM_202571.1
ABC transporter C family member 3 mRNA	825	1833	73%	NM_180244.1
ABC transporter C family member 3 mRNA	825	1893	73%	NM_112147.2
glutamate-1-semialdehyde 2,1-aminomutase 2 mRNA	801	1042	77%	NM_114732.4
pentatricopeptide repeat-containing protein EMB2745 mRNA	794	894	69%	NM_123333.1
pentatricopeptide repeat-containing protein mRNA	783	833	70%	NM_121692.1
glutamate-1-semialdehyde-2,1-aminomutase mRNA	722	984	75%	NM_125752.3
RNA recognition motif-containing protein mRNA	702	1637	71%	NM_124949.2
multidrug resistance-associated protein 8 mRNA	688	1591	71%	NM_112148.3
ABC transporter C family member 7 mRNA	670	1653	70%	NM_112149.3
tubulin alpha-2 chain mRNA	639	1355	81%	NM_103889.3
golgi nucleotide sugar transporter 3 mRNA	628	628	75%	NM_106283.2
tubulin alpha-4 chain mRNA	614	1321	81%	NM_100360.3
polyamine uptake transporter 5 mRNA	609	774	75%	NM_112845.3
vacuolar-sorting receptor 3 mRNA	580	1291	78%	NM_179624.1
vacuolar-sorting receptor 3 mRNA	580	1291	78%	NM_127038.1
tubulin alpha-6 chain mRNA	578	1325	79%	NM_117582.3

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
cuticular wax biosynthesis protein mRNA	571	2170	78%	NM_001125641.1
cuticular wax biosynthesis protein mRNA	571	2170	78%	NM_119571.4
vacuolar-sorting receptor 4 mRNA	562	1311	78%	NM_179623.2
vacuolar-sorting receptor 4 mRNA	562	1311	78%	NM_127036.4
glutamyl-tRNA reductase 1 mRNA	535	778	71%	NM_104609.3
protein TORNADO 1 mRNA	526	1269	69%	NM_124936.2
glutamyl-tRNA reductase 2 mRNA	506	727	70%	NM_100868.2
tubulin alpha-3 mRNA	489	963	77%	NM_121982.3
tubulin alpha-5 mRNA	470	991	76%	NM_121983.3
tubulin alpha-1 chain mRNA	466	737	77%	NM_105148.3
pentatricopeptide repeat-containing protein mRNA	455	514	68%	NM_127124.1
fructose-bisphosphate aldolase 5 mRNA	452	452	75%	NM_118786.3
fructose-bisphosphate aldolase 5 mRNA	452	452	75%	NM_001036644.2
translation initiation factor 3 subunit B mRNA	446	1371	79%	NM_122646.3
translation initiation factor 3 subunit B mRNA	446	1370	79%	NM_001036877.2
tubulin alpha-6 chain mRNA	441	1266	81%	NM_179057.1
eukaryotic translation initiation factor 3B-2 mRNA	435	1229	78%	NM_122479.3
seed storage transportation protein MAG2 mRNA	434	434	66%	NM_114638.2
serine acetyltransferase 2;2 mRNA	412	412	76%	NM_112150.3
serine acetyltransferase 1 mRNA	403	448	76%	NM_104470.2
pentatricopeptide repeat-containing protein mRNA	401	401	67%	NM_127679.3
fructose-bisphosphate aldolase 7 mRNA	394	529	73%	NM_118785.3
pentatricopeptide repeat-containing protein mRNA	390	494	68%	NM_127130.2
Cam interacting protein 111 mRNA	390	931	80%	NM_115528.2
methylcrotonoyl-CoA carboxylase beta chain mRNA	383	972	80%	NM_119564.4
zinc Finger RING C3H2C3-type protein mRNA	381	554	72%	NM_104844.2
BTB/POZ domain-containing protein mRNA	378	609	76%	NM_104452.3
C2 domain-containing protein mRNA	378	529	72%	NM_203207.1

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
C2 domain-containing protein mRNA	378	529	72%	NM_124935.2
C2 domain-containing protein mRNA	378	529	72%	NM_203206.1
peptidyl-prolyl cis-trans isomerase CYP40 mRNA	374	666	77%	NM_127141.3
zinc metalloprotease pitrilysin subfamily A mRNA	365	2113	76%	NM_112804.4
WD40 domain-containing protein mRNA	363	417	70%	NM_112781.2
poly(A) binding protein 8 mRNA	358	806	72%	NM_103863.2
beta-galactosidase 15 mRNA	352	352	76%	NM_148500.1
fructose-bisphosphate aldolase 6 mRNA	352	484	72%	NM_129203.2
presequence protease 2 mRNA	352	2139	76%	NM_180631.1
presequence protease 2 mRNA	352	2139	76%	NM_180630.1
presequence protease 2 mRNA	352	2139	76%	NM_103851.3
exostosin family protein mRNA	338	665	74%	NM_103149.3
chromodomain remodeling complex protein CHC1 mRNA	334	334	67%	NM_121421.3
pentatricopeptide repeat-containing protein mRNA	331	331	68%	NM_112383.1
F-box/kelch-repeat protein mRNA	327	327	67%	NM_102054.3
somatic embryogenesis receptor kinase 4 mRNA	325	817	79%	NM_126955.4
poly(A) binding protein 4 mRNA	324	555	71%	NM_127899.3
vacuolar-sorting receptor 1 mRNA	324	429	70%	NM_115145.1
40S ribosomal protein S16-1 mRNA	320	320	77%	NM_126785.1
40S ribosomal protein S16-3 mRNA	311	311	76%	NM_121843.2
somatic embryogenesis receptor kinase 5 mRNA	311	724	79%	NM_126956.3
fructose-bisphosphate aldolase mRNA	309	449	71%	NM_115153.3
transducin/WD40 domain-containing protein mRNA	307	361	68%	NM_119603.2
translation initiation factor SUI1 family protein mRNA	306	306	79%	NM_180861.2
translation initiation factor SUI1 family protein mRNA	306	306	79%	NM_124876.4
30S ribosomal protein S10 mRNA	306	306	81%	NM_112151.3
beta-galactosidase 7 mRNA	304	304	73%	NM_122078.4
uncharacterized protein mRNA	302	366	69%	NM_103948.2

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
root phototropism protein 3 mRNA	300	793	69%	NM_001126017.1
root phototropism protein 3 mRNA	300	786	69%	NM_125829.3
putative caffeoyl-CoA O-methyltransferase mRNA	298	638	83%	NM_119566.4
40S ribosomal protein S16-2 mRNA	298	298	76%	NM_111294.3
putative caffeoyl-CoA O-methyltransferase mRNA	298	298	83%	NM_179160.1
RNA-binding CRS1 / YhbY (CRM) domain protein mRNA	293	386	76%	NM_117376.4
Inositol monophosphatase family protein mRNA	293	293	76%	NM_125834.2
CDPK-related kinase mRNA	286	928	81%	NM_114913.3
S-adenosyl-L-methionine-dependent methyltransferase-like protein mRNA	286	286	71%	NM_112189.3
vacuolar sorting receptor 6 mRNA	284	455	74%	NM_102827.1
uncharacterized protein mRNA	279	352	77%	NM_114832.3
3-deoxy-d-arabino-heptulosonate 7-phosphate synthase mRNA	279	930	82%	NM_001085018.1
uncharacterized protein mRNA	279	352	77%	NM_001084796.1
3-deoxy-d-arabino-heptulosonate 7-phosphate synthase mRNA	279	1122	82%	NM_119505.2
beta-ureidopropionase mRNA	277	836	77%	NM_125833.4
L-type lectin-domain containing receptor kinase IX.1 mRNA	275	876	68%	NM_121091.1
WD40 domain-containing protein mRNA	275	275	67%	NM_128206.2
leucine-rich receptor-like protein kinase mRNA	273	327	67%	NM_124986.3
Sec14p-like phosphatidylinositol transfer family protein mRNA	266	812	82%	NM_179485.1
Sec14p-like phosphatidylinositol transfer family protein mRNA	266	812	82%	NM_179484.2
Sec14p-like phosphatidylinositol transfer family protein mRNA	266	812	82%	NM_104445.2

[z]Description of annotated gene function provided by NCBI.

[y]Calculated from the sum of the match rewards and mismatch/open gap penalties for each segment.

[x]Sum of alignment scores of all segments.

[w]Highest percent identity for a set of aligned segments.

Addendum 7. Results obtained from the BLASTn utility customized for plant genomes provided by NCBI. The genome sequence included in the 1.5 LOD interval of the *parth5.1* QTL (sequence between molecular markers UW001903 and SSR13409) was used as a query. Only the 100 matches with the highest alignment scores are reported here with matches of high interest presented first and in bold type. Matches were of high interest if they were found to associate with matches identified in other QTL regions.

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
DELLA protein GAI mRNA	**731**	**882**	**74%**	**NM_101361.2**
DELLA protein RGA mRNA	**731**	**879**	**74%**	**NM_126218.2**
auxin transport protein BIG mRNA	2419	6014	72%	NM_111093.2
sec23/sec24-like transport protein mRNA	1287	1678	76%	NM_116411.3
transducin/WD40 repeat-like superfamily protein mRNA	1173	1173	76%	NM_100933.2
putative alpha,alpha-trehalose-phosphate synthase [UDP-forming] 9 mRNA	1101	1221	73%	NM_102235.2
putative alpha,alpha-trehalose-phosphate synthase [UDP-forming] 10 mRNA	1068	1387	72%	NM_104705.3
embryo defective 2765 mRNA	1034	3150	80%	NM_179966.1
3-ketoacyl-CoA synthase 5 mRNA	964	964	75%	NM_102356.3
putative alpha,alpha-trehalose-phosphate synthase [UDP-forming] 8 mRNA	962	1257	71%	NM_105697.3
F-box/kelch-repeat protein mRNA	953	953	78%	NM_121575.4
phosphoinositide 4-kinase gamma 7 mRNA	924	1069	76%	NM_126434.3
3-ketoacyl-CoA synthase 6 mRNA	895	895	73%	NM_105524.2
putative phosphatidylinositol 4-kinase type 2-beta mRNA	877	1003	74%	NM_102391.2
phosphatidylinositol 4-kinase gamma 6 mRNA	868	868	75%	NM_101234.3
pentatricopeptide repeat-containing protein mRNA	841	934	70%	NM_127776.2
pentatricopeptide repeat-containing protein mRNA	832	832	69%	NM_105567.1
transducin/WD40 repeat-like superfamily protein mRNA	809	809	78%	NM_129939.5
ABC transporter B family member 19 mRNA	801	2735	75%	NM_113807.2
ABC transporter B family member 15 mRNA	762	1573	73%	NM_113754.2
beta-1,4-N-acetylglucosaminyltransferase family protein mRNA	726	823	77%	NM_101170.3
pentatricopeptide repeat-containing protein mRNA	724	724	70%	NM_001036910.1
COMPASS-like H3K4 histone methylation complex component mRNA	691	806	74%	NM_113000.5

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
pyruvate kinase mRNA	691	1239	79%	NM_120944.2
pyruvate kinase mRNA	652	1208	78%	NM_125763.2
pentatricopeptide repeat-containing protein mRNA	645	645	68%	NM_124421.1
pentatricopeptide repeat-containing protein mRNA	636	636	73%	NM_105775.1
putative trehalose phosphatase/synthase 5 mRNA	632	806	69%	NM_117886.2
RNA editing factor OTP85 mRNA	628	628	69%	NM_126350.2
3-ketoacyl-CoA synthase 6 mRNA	625	902	73%	NM_179530.1
endosomal targeting BRO1-like domain-containing protein mRNA	619	1588	76%	NM_101381.2
leucine-rich repeat protein kinase-like protein mRNA	609	609	69%	NM_113765.2
auxin efflux carrier family protein mRNA	607	607	71%	NM_105778.1
U-box domain-containing protein 14 mRNA	605	678	73%	NM_115342.4
putative glycosyl transferase mRNA	605	605	72%	NM_111899.2
beta-1,4-N-acetylglucosaminyltransferase like protein mRNA	605	660	73%	NM_105458.2
auxin efflux carrier component 3 mRNA	585	831	70%	NM_105762.2
RING/U-box domain-containing protein mRNA	567	1012	71%	NM_120528.3
phosphoinositide 4-kinase gamma 7 mRNA	560	754	76%	NM_201684.1
leucine-rich repeat-containing protein kinase mRNA	556	1402	69%	NM_102342.1
SNF2, helicase and F-box domain-containing protein mRNA	554	1425	70%	NM_148874.4
uncharacterized protein mRNA	549	549	87%	NM_130279.2
ABC transporter B family member 17 mRNA	545	1433	69%	NM_113758.1
galactose oxidase/kelch repeat superfamily protein mRNA	540	540	71%	NM_101299.3
ABC transporter B family member 22 mRNA	538	1370	69%	NM_148757.1
Retinoblastoma-related protein 1 mRNA	536	1600	74%	NM_112064.4
auxin efflux carrier component 7 mRNA	524	671	72%	NM_001084115.1
auxin efflux carrier component 7 mRNA	524	740	72%	NM_102156.1
auxin efflux carrier component 7 mRNA	524	671	72%	NM_179369.1
autoinhibited Ca2+/ATPase II mRNA	524	1537	69%	NM_101192.2
pentatricopeptide repeat-containing protein mRNA	518	597	70%	NM_104760.2

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
ubiquitin-40S ribosomal protein S27a-3 mRNA	518	518	83%	NM_116090.2
beta-1,4-N-acetylglucosaminyltransferase family protein mRNA	515	515	72%	NM_113670.3
ABC transporter B family member 18 mRNA	511	1138	69%	NM_113759.1
putative serine/threonine protein kinase mRNA	509	1062	80%	NM_105425.4
aspartyl protease family protein mRNA	509	565	71%	NM_102362.4
ATPase E1-E2 type protein/haloacid dehalogenase-like hydrolase mRNA	508	1111	69%	NM_113459.1
beta-1,4-N-acetylglucosaminyltransferase family protein mRNA	508	508	72%	NM_121452.3
auxin efflux carrier component 4 mRNA	506	1004	71%	NM_126203.2
auxin efflux carrier component 4 mRNA	506	1004	71%	NM_179592.1
F-box/kelch-repeat protein SKIP11 mRNA	504	504	71%	NM_126342.3
protein EMBRYO DEFECTIVE 1220 mRNA	504	736	73%	NM_104707.3
F-box/kelch-repeat protein SKIP11 mRNA	504	504	71%	NM_001035883.1
F-box/kelch-repeat protein SKIP11 mRNA	504	504	71%	NM_001035882.1
E3 ubiquitin-protein ligase KEG mRNA	497	2654	69%	NM_121356.2
reversably-glycosylated protein 5 mRNA	486	486	73%	NM_180500.2
reversably-glycosylated protein 5 mRNA	486	486	73%	NM_121657.1
Small GTP-binding protein mRNA	484	983	75%	NM_123353.3
phosphoenolpyruvate carboxylase 4 mRNA	484	1710	73%	NM_105548.4
beta-1,4-N-acetylglucosaminyltransferase family protein mRNA	477	477	70%	NM_111028.2
phosphatidylinositol-4-phosphate 5-kinase 6 mRNA	475	1267	78%	NM_111675.5
putative galacturonosyltransferase-like 9 mRNA	471	704	73%	NM_001124104.1
putative galacturonosyltransferase-like 9 mRNA	471	704	73%	NM_105677.2
Sphingoid long-chain bases kinase 1 mRNA	464	1053	71%	NM_180734.2
Sphingoid long-chain bases kinase 1 mRNA	464	1053	71%	NM_122252.3
U-box domain-containing protein 45 mRNA	464	575	69%	NM_102556.4
SRP72 RNA-binding domain-containing protein mRNA	452	891	72%	NM_105436.3
protein CHUP1 mRNA	448	1507	72%	NM_113468.4
Calcium dependent protein kinase 1 mRNA	444	1091	77%	NM_120569.2

Description[z]	Max[y]	Total[x]	Ident[w]	Accession
Calcium-dependent protein kinase 2 mRNA	441	1154	78%	NM_111902.1
cleavage and polyadenylation specificity factor subunit 73-I mRNA	437	2317	75%	NM_104782.3
cleavage and polyadenylation specificity factor subunit 73-I mRNA	437	2317	75%	NM_001036138.1
cleavage and polyadenylation specificity factor subunit 73-I mRNA	437	2317	75%	NM_179504.1
photosystem I light harvesting complex protein mRNA	432	536	79%	NM_116012.4
protein kinase superfamily protein mRNA	432	432	73%	NM_102350.2
3-ketoacyl-CoA synthase 9 mRNA	432	432	69%	NM_127184.2
putative galacturonosyltransferase-like 8 mRNA	432	646	71%	NM_102263.2
putative galacturonosyltransferase-like 3 mRNA	432	432	72%	NM_101196.2
ATP-dependent zinc metalloprotease FTSH 6 mRNA	428	1119	77%	NM_121529.2
leucine-rich repeat protein kinase family protein mRNA	428	668	76%	NM_102307.1
pentatricopeptide repeat-containing protein mRNA	421	421	68%	NM_111284.2
pentatricopeptide repeat-containing protein mRNA	421	421	68%	NM_001035553.1
leucine-rich repeat transmembrane protein kinase-like protein mRNA	419	599	67%	NM_126182.3
potassium transporter 8 mRNA	419	2139	69%	NM_121492.1
leucine-rich repeat protein kinase family protein mRNA	417	417	67%	NM_102481.3
ribose 5-phosphate isomerase A mRNA	414	810	74%	NM_126190.1
nitrate transporter 1:2 mRNA	412	703	72%	NM_105653.4
SRP72 RNA-binding domain protein mRNA	407	735	71%	NM_105433.2
transducin/WD-40 repeat-containing protein mRNA	407	407	71%	NM_102297.2
60S ribosomal protein L27a-3 mRNA	405	405	80%	NM_105728.3

[z]Description of annotated gene function provided by NCBI.

[y]Calculated from the sum of the match rewards and mismatch/open gap penalties for each segment.

[x]Sum of alignment scores of all segments.

[w]Highest percent identity for a set of aligned segments.

Addendum 8. Phenotypic data used in all analyses of the 2A×Gy8 $F_{2:3}$ cucumber population. The seed size and seed weight traits were collected from the F_2 generation. Parthenocarpic fruit set data was collected from the F_3 generation and presented as a mean of those values for each F_3 family.

Family	Parthenocarpic Fruit Set[z]	Seed Size (cm^2)[y]	Seed Weight (g)[x]
1	4.36	8.55	1.1
2	3.45	8.80	1.3
3	5.09	8.55	1.6
4	5.36	8.08	1.4
5	2.55	7.74	1.1
6	4.36	7.98	1.0
7	2.27	6.15	1.1
8	1.91	6.24	0.9
9	3.27	8.40	1.2
10	4.55	8.55	1.4
11	2.00	8.55	1.5
12	2.50	6.40	1.1
13	4.09	8.36	1.3
14	4.00	8.55	1.3
15	4.00	8.55	1.3
16	2.18	7.56	1.2
17	2.45	9.20	1.5
18	2.80	6.72	1.1
19	1.82	5.78	0.8
20	2.91	5.10	0.9
21	4.45	7.20	1.4
22	1.91	8.74	1.6
23	4.00	8.36	1.1
24	2.09	7.74	1.1
25	4.91	9.70	1.6
26	3.09	9.80	1.6
27	2.82	8.10	1.3
28	2.91	8.74	1.5
29	5.18	9.20	1.7
30	2.27	6.40	1.0
31	2.73	5.40	1.0
33	3.00	4.95	0.7
34	5.73	7.14	1.4

Family	Parthenocarpic Fruit Set[z]	Seed Size (cm^2)[y]	Seed Weight (g)[x]
35	4.00	8.93	1.7
36	3.50	8.36	1.3
37	3.55	9.66	1.4
38	4.20	6.97	1.2
39	3.18	6.97	1.4
40	2.18	5.60	0.9
41	4.45	6.72	1.1
42	3.64	5.70	1.0
44	4.91	7.98	1.5
45	3.73	7.79	1.2
46	4.27	9.00	1.4
47	2.73	8.74	1.4
49	4.64	5.85	1.1
50	2.55	8.36	1.4
51	3.36	7.31	1.2
52	3.89	6.72	1.3
53	2.18	5.55	0.9
54	2.09	6.00	1.1
55	4.64	5.85	1.0
56	6.60	8.74	1.4
57	4.45	8.55	1.4
58	3.45	6.00	1.0
59	5.09	6.88	1.4
60	3.45	4.48	0.9
61	3.55	5.04	1.1
62	4.90	7.14	1.2
63	3.18	6.63	1.1
64	5.27	8.20	1.3
65	2.91	7.98	1.2
66	2.50	7.48	1.4
67	2.09	8.46	1.5
68	4.82	6.80	1.3
69	2.50	6.40	0.9
72	2.55	6.56	1.2
73	2.36	6.72	1.4
74	5.45	6.88	1.2
75	4.45	7.48	1.5
76	2.55	6.40	1.2

Family	Parthenocarpic Fruit Set[z]	Seed Size (cm^2)[y]	Seed Weight (g)[x]
77	1.80	5.10	0.8
78	3.91	6.97	1.2
79	2.10	6.97	1.4
80	2.73	5.18	1.1
81	3.73	7.92	1.5
82	3.40	7.31	1.5
83	1.73	5.60	1.1
84	4.27	5.85	1.2
85	3.82	8.55	1.5
86	4.10	6.40	1.2
87	3.00	7.56	1.3
88	3.55	6.97	1.1
89	1.91	7.38	1.2
90	4.70	7.23	1.4
91	4.30	6.56	1.2
92	2.45	5.10	0.9
93	5.00	5.85	1.1
94	5.64	7.74	1.6
95	2.18	4.20	0.6
97	4.73	6.40	1.2
98	3.09	6.00	1.0
99	3.73	6.97	1.3
100	3.36	7.79	1.4
101	2.91	8.10	1.7
102	3.64	6.08	1.1
103	3.91	5.60	0.9
104	3.45	8.60	1.5
105	2.70	7.48	1.3
107	4.82	6.40	1.2
108	3.27	7.20	1.4
109	4.09	6.88	1.4
110	3.82	7.82	1.5
111	2.27	8.17	1.4
113	4.73	7.98	1.6
114	2.60	6.08	1.2
115	2.73	7.02	1.4
116	2.18	8.55	1.5
117	2.64	7.31	1.3

Family	Parthenocarpic Fruit Set[z]	Seed Size (cm^2)[y]	Seed Weight (g)[x]
118	5.55	8.80	1.5
119	3.00	9.00	1.5
120	3.27	6.24	1.2
121	3.45	9.00	1.7
122	5.00	7.41	1.1
123	3.55	8.10	1.2
124	2.27	7.92	1.4
125	4.55	9.00	1.5
126	4.36	6.24	1.2
127	1.18	7.14	1.2
128	6.27	6.30	1.2
129	3.36	7.79	1.5
131	1.82	5.70	1.0
133	3.73	6.15	1.2
134	4.09	7.14	1.3
135	3.82	9.80	2.0
136	1.20	4.80	0.9
137	4.00	7.20	1.2
138	6.00	6.97	1.3
139	3.45	4.48	0.8
140	5.50	5.10	1.0
141	2.55	7.14	1.3
142	2.82	6.40	1.2
143	3.09	8.10	1.6
144	2.55	8.17	1.7
145	1.73	8.80	1.2
146	3.00	5.10	0.8
147	3.45	6.24	1.1
148	4.00	6.80	1.4
149	3.55	7.48	1.6
150	3.36	8.40	1.6
151	2.82	6.97	1.3
152	4.00	8.55	1.6
153	2.50	7.56	1.4
154	5.27	7.38	1.2
155	1.60	5.44	0.7
156	2.50	6.00	1.1
157	2.64	4.90	0.9

Family	Parthenocarpic Fruit Set[z]	Seed Size (cm^2)[y]	Seed Weight (g)[x]
158	2.20	6.46	1.2
161	3.00	9.00	1.3
162	2.00	5.60	NA
164	4.27	7.38	1.3
165	2.09	6.40	1.3
167	4.45	6.88	1.3
168	3.00	9.00	NA
169	4.70	7.20	1.1
170	4.55	7.98	1.1
171	1.70	5.25	1.1
172	4.40	5.44	1.1
173	3.20	5.10	0.9
174	5.91	8.28	1.7
175	3.00	6.40	1.1
176	3.90	6.97	1.3
177	1.80	4.80	0.9
178	5.64	7.20	1.4
179	2.64	7.79	1.3
180	3.73	6.45	1.4
181	2.18	7.31	1.6
182	2.73	8.80	1.6
184	1.00	5.25	0.7
185	4.09	6.24	1.0
186	3.55	5.55	1.1
187	4.55	5.92	1.2
188	2.64	7.74	1.7
189	2.64	6.29	1.2
190	4.90	5.76	1.1
191	1.50	6.40	1.0
192	2.36	6.72	1.1
193	3.36	7.82	1.5
196	4.00	7.31	1.2
198	4.73	8.28	1.5
199	5.55	7.92	1.4
200	5.27	7.02	1.5
201	4.82	7.74	1.4
202	5.73	8.10	1.7
203	2.50	6.24	1.2

Family	Parthenocarpic Fruit Set[z]	Seed Size (cm^2)[y]	Seed Weight (g)[x]
205	3.00	6.12	0.8
206	3.00	6.46	1.1
207	4.18	8.93	1.7
208	2.09	6.08	0.9
209	2.55	7.02	1.5
210	2.55	6.24	1.2
211	3.09	6.97	0.9
212	3.30	5.44	1.0
213	3.55	5.25	0.9
214	3.60	6.97	1.3
215	3.82	6.15	1.1
216	5.73	7.65	1.5
217	4.89	6.97	1.2
218	3.73	7.38	1.3
219	1.82	7.14	1.5
220	3.73	7.65	1.5
221	3.18	7.14	1.4
222	4.00	6.29	1.1
223	3.91	6.72	1.1
224	2.64	7.98	1.5
2A	5.52	5.18	0.9
Gy8	3.08	8.05	1.4
2A×Gy8 F1	2.93	NA	NA

[z]Parthenocarpic fruit set was measured as the number of parthenocarpic fruits initiated on each plant. Values presented here are the means of 11 F_3 individuals.

[y]Seed size was scored as the mean length (cm) multiplied by the mean width (cm) of five seeds from a single fruit for each plant. Mean length and width measurements were taken from the longest and widest dimension of five healthy and fully developed seeds.

[x]Seed weight was scored as the weight in grams of 50 healthy and fully developed seeds from a single fruit.

Addendum 9. Correlation coefficients calculated from comparisons of parthenocarpic fruit set, seed size, and seed weight traits in a 2AxGy8 $F_{2:3}$ cucumber population.

	Parthenocarpic Fruit Set	Seed Size	Seed Weight
Parthenocarpic Fruit Set		0.22**	0.27***
Seed Size	0.23**		0.79***
Seed Weight	0.27***	0.79***	

***Calculated values were found to be significant at alpha = 0.01.

** Calculated values were found to be significant at alpha = 0.05.

Addendum 10. Alignment of the predicted protein sequences of the candidate gene BRI1, obtained from the parental lines '2A' and 'Gy8'. Protein sequences were predicted with assembled sequence data obtained from whole genome re-sequencing of the parental lines. The predicted BRI1 protein from 'Gy14' was constructed from sequence data extracted from the Gy14 Draft Genome Assembly Version 1.0 and is included as a reference (Yang et al., 2012). Protein prediction was performed with the FGENESH utility provided by Softberry (Solovyev et al., 2006). The gap in sequence data observed for 'Gy8' is due to a gap between contigs of the 'Gy8' assembled re-sequencing data. An asterisk marks a potential polymorphism between '2A' and the other sequences.

```
2A      MIPFFPSSSNSFLTFFFFFVSLTFLSFSVSSVTPSSSHGDTQKLVSFKASLPNPTLLQNW
GY14    MIPFFPSSSNSFLTFFFFFVSLTFLSFSVSSVTPSSSHGDTQKLVSFKASLPNPTLLQNW
GY8     MIPFFPSSSNSFLTFFFFFVSLTFLSFSVSSVTPSSSHGDTQKLVSFKASLPNPTLLQNW

2A      LSNADPCSFSGITCKETRVSAIDLSFLSLSSNFSHVFPLLAALDHLESLSLKSTNLTGSI
GY14    LSNADPCSFSGITCKETRVSAIDLSFLSLSSNFSHVFPLLAALDHLESLSLKSTNLTGSI
GY8     LSNADPCSFSGITCKETRVSAIDLSFLSLSSNFSHVFPLLAALDHLESLSLKSTNLTGSI

2A      SLPSGFKCSPLLASVDLSLNGLFGSVSDVSNLGFCSNVKSLNLSFNAFDFPLKDSAPGLK
GY14    SLPSGFKCSPLLASVDLSLNGLFGSVSDVSNLGFCSNVKSLNLSFNAFDFPLKDSAPGLK
GY8     SLPSGFKCSPLLASVDLSLNGLFGSVSDVSNLGFCSNVKSLNLSFNAFDFPLKDSAPGLK

2A      LDLQVLDLSSNRIVGSKLVPWIFSGGCGSLQHLALKGNKISGEINLSSCNKLEHLDISGN
GY14    LDLQVLDLSSNRIVGSKLVPWIFSGGCGSLQHLALKGNKISGEINLSSCNKLEHLDISGN
GY8     LDLQVLDLSSNRIVGSKLVPWIFSGGCGSLQHLALKGNKISGEINLSSCNKLEHLDISGN

2A      NFSVGIPSLGDCSVLEHFDISGNKFTGDVGHALSSCQQLTFLNLSSNQFGGPIPSFASSN
GY14    NFSVGIPSLGDCSVLEHFDISGNKFTGDVGHALSSCQQLTFLNLSSNQFGGPIPSFASSN
GY8     NFSVGIPSLGDCSVLEHFDISGNKFTGDVGHALSSCQQLTFLNLSSNQFGGPIPSFASSN

2A      LWFLSLANNDFQGEIPVSIADLCSSLVELDLSSNSLIGAVPTALGSCFSLQTLDISKNNL
GY14    LWFLSLANNDFQGEIPVSIADLCSSLVELDLSSNSLIGAVPTALGSCFSLQTLDISKNNL
GY8     LWFLSLANNDFQGEIPVSIADLCSSLVELDLSSNSLIGAVPTALGSCFSLQTLDISKNNL

2A      TGELPIAVFAKMSSLKKLSVSDNKFFGVLSDSLSQLAILNSLDLSSNNFSGSIPAGLCED
GY14    TGELPIAVFAKMSSLKKLSVSDNKFFGVLSDSLSQLAILNSLDLSSNNFSGSIPAGLCED
GY8     TGELPIAVFAKMSSLKKLSVSDNKFFGVLSDSLSQLAILNSLDLSSNNFSGSIPAGLCED

2A      PSNNLKELFLQNNWLTGRIPASISNCTQLVSLDLSFNFLSGTIPSSLGSLSKLKNLIMWL
GY14    PSNNLKELFLQNNWLTGRIPASISNCTQLVSLDLSFNFLSGTIPSSLGSLSKLKNLIMWL
GY8     PSNNLKELFLQNNWLTGRIPASISNCTQLVSLDLSFNFLSGTIPSSLGSLSKLKNLIMWL

2A      NQLEGEIPSDFSNFQGLENLILDFNELTGTIPSGLSNCTNLNWISLSNNRLKGEIPAWIG
GY14    NQLEGEIPSDFSNFQGLENLILDFNELTGTIPSGLSNCTNLNWISLSNNRLKGEIPAWIG
GY8     NQLEGEIPSDFSNFQGLENLILDFNELTGTIPSGLSNCTNLNWISLSNNRLKGEIPAWIG
```

Addendum 10 continued.

```
2A      SLPNLAILKLSNNSFYGRIPKELGDCRSLIWLDLNTNLLNGTIPPELFRQSGNIAVNFIT
GY14    SLPNLAILKLSNNSFYGRIPKELGDCRSLIWLDLNTNLLNGTIPPELFRQSGNIAVNFIT
GY8     SLPNLAILKLSNNSFYGRIPKELGDCRSLIWLDLNTNLLNGTIPPELFRQSGNIAVNFIT

2A      GKSYAYIKNDGSKQCHGAGNLLEFAGIRQEQVNRISSKSPCNFTRVYKGMIQPTFNHNGS
GY14    GKSYAYIKNDGSKQCHGAGNLLEFAGIRQEQVNRISSKSPCNFTRVYKGMIQPTFNHNGS
GY8     GKSYAYIKNDGSKQCHGAGNLLEFAGIRQEQVNRISSKSPCNFTRVYKGMIQPTFNHNGS

2A      MIFLDLSHNMLTGSIPKDIGSTNYLYILDLGHNSLSGPIPQELGDLTKLNILDLSGNELE
GY14    MIFLDLSHNMLTGSIPKDIGSTNYLYILDLGHNSLSGPIPQELGDLTKLNILDLSGNELE
GY8     MIFLDLSHNMLTGSIPKDIGSTNYLYILDLGHNSLSGPIPQELGDLTKLNILDLSGNELE

2A      GSIPLSLTGLSSLMEIDLSNNHLNGSIPESAQFETFPASGFANNSGLCGYPLPPCVVDSA
GY14    GSIPLSLTGLSSLMEIDLSNNHLNGSIPESAQFETFPASGFANNSGLCGYPLPPCVVDSA
GY8     GSIPLSLTGLSSLMEIDLSNNHLNGSIPESAQFETFPASGFANNSGLCGYPLPPCVVDSA

2A      GNANSQHQRSHRKQASLAGSVAMGLLFSLFCIFGLIIVVIEMRKRRKKKDSALGSYVESH
GY14    GNANSQHQRSHRKQASLAGSVAMGLLFSLFCIFGLIIVVIEMRKRRKKKDSALDSYVESH
GY8     GNANSQHQRSHRKQASLAGSVAMGLLFSLFCIFGLIIVVIEMRKRRKKKDSALDSYVESH
                                                                      *
2A      SQSGTTTAVNWKLTGAREALSINLATFEKPLRKLTFADLLEATNGFHNDSLIGSGGFGDV
GY14    SQSGTTTAVNWKLTGAREALSINLATFEKPLRKLTFADLLEATNGFHNDSLIGSGGFGDV
GY8     SQSGTTTAVNWKLT---------------------------------------------------------------- ----GGFGDV

2A      YKAQLKDGSTVAIKKLIHVSGQGDREFTAEMETIGKIKHRNLVPLLGYCKVGEERLLVYE
GY14    YKAQLKDGSTVAIKKLIHVSGQGDREFTAEMETIGKIKHRNLVPLLGYCKVGEERLLVYE
GY8     YKAQLKDGSTVAIKKLIHVSGQGDREFTAEMETIGKIKHRNLVPLLGYCKVGEERLLVYE

2A      YMKYGSLEDVLHDQKKGGIKLNWSARRKIAIGAARGLAFLHHNCIPHIIHRDMKSSNVLL
GY14    YMKYGSLEDVLHDQKKGGIKLNWSARRKIAIGAARGLAFLHHNCIPHIIHRDMKSSNVLL
GY8     YMKYGSLEDVLHDQKKGGIKLNWSARRKIAIGAARGLAFLHHNCIPHIIHRDMKSSNVLL

2A      DENLEARVSDFGMARLMSAMDTHLSVSTLAGTPGYVPPEYYQSFRCSTKGDVYSYGVVML
GY14    DENLEARVSDFGMARLMSAMDTHLSVSTLAGTPGYVPPEYYQSFRCSTKGDVYSYGVVML
GY8     DENLEARVSDFGMARLMSAMDTHLSVSTLAGTPGYVPPEYYQSFRCSTKGDVYSYGVVML

2A      ELLTGKRPTDSADFGDNNLVGWVKQHVKLDPIDVFDPELIKEDPSLKIELLEHLKVAVAC
GY14    ELLTGKRPTDSADFGDNNLVGWVKQHVKLDPIDVFDPELIKEDPSLKIELLEHLKVAVAC
GY8     ELLTGKRPTDSADFGDNNLVGWVKQHVKLDPIDVFDPELIKEDPSLKIELLEHLKVAVAC

2A      LDDRSWRRPTMIQVMTMFKEIQAGSGMDSHSTIGTDNGGFSVDMVDMSLKEVPEPEGK
GY14    LDDRSWRRPTMIQVMTMFKEIQAGSGMDSHSTIGTDNGGFSVDMVDMSLKEVPEPEGK
GY8     LDDRSWRRPTMIQVMTMFKEIQAGSGMDSHSTIGTDNGGFSVDMVDMSLKEVPEPEGK
```

Addendum 11. Alignment of the predicted protein sequences of the candidate gene BAK1, obtained from the parental lines '2A' and 'Gy8'. Protein sequences were predicted with assembled sequence data obtained from whole genome re-sequencing of the parental lines. The predicted BAK1 protein from 'Gy14' was constructed from sequence data extracted from the Gy14 Draft Genome Assembly Version 1.0 and is included as a reference (Yang et al., 2012). Protein prediction was performed with the FGENESH utility provided by Softberry (Solovyev et al., 2006). The gap in sequence data observed for '2A' and 'Gy8' is due to gaps between contigs of the assembled re-sequencing data. The mismatch of sequence flanking the gap between contigs of '2A' is a result of an overhanging base pair attached to the edge of the first contig which resulted in a change to the predicted protein around this gap. An asterisk marks a potential polymorphism between '2A' and the other sequences.

```
2A      MRRKCLGWSLSRHFPRCSAKAFLTAFGQLVLPFGSDVDDHLMEMEQYKVLALGFVSLILL
GY14    MRRKCLGWSLSRHFPRCSAKAFLTAFGQLVLPFGSDVDDHLMEMEQYKVLALGFVSLILL
GY8     MRRKCLGWSLSRHFPRCSAKAFLTAFGQLVLPFGSDVDDHLMEMEQYKVLALGFVSLILL

2A      VRPLWLVSANMEGDALHSLRTSLQDPNNVLQSWDPTLVNPCTWFHVTCNNDNSVIRVDLG
GY14    VRPLWLVSANMEGDALHSLRTSLQDPNNVLQSWDPTLVNPCTWFHVTCNNDNSVIRVDLG
GY8     VRPLWLVSANMEGDALHSLRTSLQDPNNVLQSWDPTLVNPCTWFHVTCNNDNSVIRVDLG

2A      NAALSGTLVPQLGLLKNLQYLELYSNNISGVIPSDLGNLTSLVSLDLYLNRFSGPIPDTL
GY14    NAALSGTLVPQLGLLKNLQYLELYSNNISGVIPSDLGNLTSLVSLDLYLNRFSGPIPDTL
GY8     NAALSGTLVPQLGLLKNLQYLELYSNNISGVIPSDLGNLTSLVSLDLYLNRFSGPIPDTL

2A      GKLSKLRFLFVYFLHCFFLECFNKDSRLNNNSLAGPIPMSLTNISSLQVLDLSNNHLSGV
GY14    GKLSKLRFLFVYFLHCFFLECFNKDSRLNNNSLAGPIPMSLTNISSLQVLDLSNNHLSGV
GY8     GKLSKLRFLFVYFLHCFFLECFNKDSRLNNNSLAGPIPMSLTNISSLQVLDLSNNHLSGV

2A      VPDNGSFSLFTPISFANNLDLCGPVTGRPCPGSPPFSPPPPFVPPPPISSPGMKMSSLVE
GY14    VPDNGSFSLFTPISFANNLDLCGPVTGRPCPGSPPFSPPPPFVPPPPISSPGMKMSSLVE
GY8     VPDNGSFSLFTPISFANNLDLCGPVTGRPCPGSPPFSPPPPFVPPPPISSPGMKMSSLVE

2A      KIQTQPSNVRMVFLSSPGKKMVRYVYPSSPCKMVDGFFFLSPELTPQSSFLTRALTMLTD
GY14    KIQTQPSNVRMVFLSSPGKKMVRYVYPSSPCKMVDGFFFLSPELTPQSSFLTRALTMLTD
GY8     KIQTQPSNVRMVFLSSPGKKMVRYVYPSSPCKMVDGFFFLSPELTPQSSFLTRALTMLTD

2A      GSVGWWHVIDNRVEGDFSSQNGGGNSATGAIAGGVAAAAALLFAAPAIAFAWWRRRKPQE
GY14    GSVGWWHVIDNRVEGDFSSQNGGGNSATGAIAGGVAAGAALLFAAPAIAFAWWRRRKPQE
GY8     GSVGWWHVIDNRVEGDFSSQNGGGNSATGAIAGGVAAGAALLFAAPAIAFAWWRRRKPQE
                                                  *
2A      VFFDVPAEEDPEVHLGQLKRFSLRELQVATDSFRRTYARWRAAVSN----------
GY14    VFFDVPAEEDPEVHLGQLKRFSLRELQVATDSFSNKNILGRGGFGKVYKGRLADGSLVAV
GY8     VFFDVPAEEDPEVHLGQLKRFSLRELQVATDSFT----------------------------------------------V
```

Addendum 11 continued.

```
2A      ---------------------------------RSRDDQHGCAPNLLRLRGFCMTPTERLLVYPYMANGSVASCLR
GY14    KRLKEERTPGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMANGSVASCLR
GY8     KRLKEERTPGGELQFQTEVEMISMAVHRNLLRLRGFCMTPTERLLVYPYMANGSVASCLR

2A      ERPPSQPPLDWRTRKRIALGSARGLSYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFG
GY14    ERPPSQPPLDWRTRKRIALGSARGLSYLHDHCDPKIIHRDVKAANILLDEEFEAVVGDFG
GY8     ----------------------------------------------------------------------------------------GDFG

2A      LAKLMDYKDTHVTTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGIMLLELITGQRAFDLAR
GY14    LAKLMDYKDTHVTTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGIMLLELITGQRAFDLAR
GY8     LAKLMDYKDTHVTTAVRGTIGHIAPEYLSTGKSSEKTDVFGYGIMLLELITGQRAFDLAR

2A      LANDDDVMLLDWVKGLLKEKKLEMLVDPDLQNNYIESEVEQLIQVALLCTQGSPMDRPKM
GY14    LANDDDVMLLDWVKGLLKEKKLEMLVDPDLQNNYIESEVEQLIQVALLCTQGSPMDRPKM
GY8     LANDDDVMLLDWVKGLLKEKKLEMLVDPDLQNNYIESEVEQLIQVALLCTQGSPMDRPKM

2A      SEVVRMLEGDGLAERWDEWQKVEILRQEIDLSPHPNSDWIVDSTENLHAVELSGPR
GY14    SEVVRMLEGDGLAERWDEWQKVEILRQEIDLSPHPNSDWIVDSTENLHAVELSGPR
GY8     SEVVRMLEGDGLAERWDEWQKVEILRQEIDLSPHPNSDWIVDSTENLHAVELSGPR
```