

A Consensus Genetic Map and Linkage Panel for Fresh-market Tomato

Prashant Bhandari

Horticultural Sciences Department, University of Florida, Gainesville, FL 32611

Reza Shekasteband

Department of Horticultural Science, North Carolina State University, Mountain Horticultural Crops Research & Extension Center, Mills River, NC 28759

Tong Geon Lee

Horticultural Sciences Department, University of Florida, Gainesville, FL 32611; Gulf Coast Research and Education Center, University of Florida, Wimauma, FL 33598; and Plant Breeders Working Group and Plant Molecular and Cellular Biology Graduate Program, University of Florida, Gainesville, FL 32611

ADDITIONAL INDEX WORDS. applied tomato genetics, breeding, plant biology, *Solanum lycopersicum*

ABSTRACT. The first consensus genetic map in fresh-market tomato (*Solanum lycopersicum*) was constructed, combining genetic recombination data from two biparental F₂ segregating populations derived from four different fresh-market tomatoes. Each F₂ population was nominated by different academic tomato breeding programs located in major fresh-market tomato-producing areas of the United States, and chromosome-wide variation in recombination rates was observed between tomato populations based on the origin of their breeding programs. A consensus map constructed using 335 common single nucleotide polymorphism (SNP) sites found in both populations spanned 737.3 cM across 12 tomato chromosomes, with chromosome 2 containing more than 40% of the total SNPs and chromosomes 4, 5, 7, and 10 together representing less than 10% of the SNPs. There was a high degree of collinearity between the genetic and physical positions of those 335 SNP markers. The integration of 6553 SNP sites that were detected in either of the two populations with 335 common sites resulted in an extended consensus genetic map. The total length of the extended map was estimated to be 1997.9 cM, which was compatible with a previous estimate for large-fruited fresh-market tomato. A linkage panel for fresh-market tomato was also established using the combined dataset of the consensus map of 335 SNP loci and 73 SNP-genotyped core fresh-market tomatoes. An empirical genetic mapping study of the tomato *brachytic* trait using the linkage panel demonstrated the value of the consensus map and linkage panel for tomato research. The allelic information in the linkage panel will serve as a basis for SNP marker implementation, such as genotyping platforms and genomic association map, in tomato.

Tomato (*Solanum lycopersicum*) is the highest value horticultural crop worldwide (Food and Agriculture Organization of the United Nations, 2016), providing micronutrients in the human diet [U.S. Department of Agriculture (USDA), 2016]. Fresh-market and processing tomatoes are the two most consumed types of tomatoes. The former is bred for direct consumption, and per capita consumption of such tomatoes in the United States alone is ≈ 9.5 kg per year (USDA, 2016).

Genetic improvements through domestication and breeding and/or plant biological research have successfully improved tomato traits, most notably disease resistance (Lukyanenko, 1991; Scott et al., 2013). However, the genetic architecture (i.e., characteristics of sequence variations responsible for traits) of complex, quantitative traits (e.g., fruit yield) in the modern

tomato (i.e., cultivated tomato) that has been bred for consumers to satisfy market demands and production systems needs to be further explored to improve breeding methods for such tomato and production systems. Specifically, in contrast to genetic variations (i.e., distinct types of DNA sequence polymorphisms) already introgressed into such tomatoes, the genetic architecture of phenotypic diversities within specific market classes of cultivated tomato is less well understood. This is presently evidenced by the fact that the analysis of such variation in tomato has often focused on differences between modern and wild tomatoes, and most of the work done to date has included a few modern breeding materials. For example, the previously characterized major loci for fruit size and shape variation [*fasciated*, *fw2.2*, *fw3.2*, *locule-number*, *ovate*, *sun* (Rodríguez et al., 2011; Tanksley, 2004)] were found to be already fixed in most large-fruited fresh-market tomato germplasm from two public fresh-market tomato breeding programs [Blanca et al., 2015; North Carolina State University (NCSU), 2021; University of Florida/Institute of Food and Agricultural Sciences (UF/IFAS), 2021]. Similarly, studies on the genetic variations currently present at a low frequency in modern germplasm (e.g., genes associated with organ architecture traits such as inflorescence architecture) were not widely used by tomato breeding programs because traits controlled by such variations were not often very high priorities for

Received for publication 19 July 2021. Accepted for publication 13 Oct. 2021.

Published online 16 December 2021.

This work is supported by the Florida Tomato Committee (T.G. Lee) and U.S. Department of Agriculture, National Institute of Food and Agriculture, Hatch project FLA-GCC-005550 (T.G. Lee). We thank Randolph G. Gardner and Samuel F. Hutton for sharing tomato material and providing knowledge of breeding history.

T.G.L. is the corresponding author. E-mail: tonggeonlee@ufl.edu.

This is an open access article distributed under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

modern tomato improvement (S.F. Hutton, personal communication). Therefore, it is necessary to develop resources for both tomato breeding and genetics, especially to understand modern germplasm in breeding programs.

Genetic maps play an important role in the characterization of genomes and contribute to trait improvement in breeding. In practice, genetic improvements in crop plants have been achieved by our knowledge of inheritance patterns of genes with known functions or DNA fragments/quantitative trait loci (QTL) of desirable traits, which have been coordinated along the physical map (i.e., the draft genome assembly or complete genome sequence of a species). Given that such inheritance can be traced in a genetic map that has been constructed based on linkage with the recombination rate, the genetic map is a solid foundation for genetic improvements in crop plants. More importantly, a consensus genetic map that depicts the position (both marker order and distance between flanking markers) of genetic markers for various plant materials from a single species can be a common framework for such applications. Genetic maps in tomato were developed using a cross between the modern tomato species (*S. lycopersicum*) and wild tomato species (e.g., *Solanum pimpinellifolium*, *Solanum pennellii*) for both tomato genetics and breeding approaches, resulting in direct impact on modern tomato improvement (Ashrafi et al., 2009; Chetelat et al., 2000; de Vicente and Tanksley, 1993; Gonda et al., 2019; Robbins et al., 2011; Sim et al., 2012a; Tanksley and Rick, 1980; Tanksley et al., 1992). Similarly, tomato consensus maps were developed from the crosses between the modern tomato species and wild tomato species (Table 1). The genetic map of fresh-market tomato that depicts the recombination rate in a cross between two fresh-market tomatoes (*S. lycopersicum* × *S. lycopersicum*), which is the most common type of initial crosses in fresh-market tomato improvement programs, was constructed very recently (Bhandari and Lee, 2021), but the consensus genetic map for this type of cross has never been constructed.

In this study, we constructed the first consensus genetic map of fresh-market tomato using whole-genome sequencing (WGS)-based SNP genotyping and two biparental F₂ segregating populations nominated by two different fresh-market tomato breeding programs. In addition, we developed a linkage panel using the combined dataset of the consensus map and 73 SNP-genotyped core fresh-market tomatoes. The core fresh-market tomatoes included three distinct fresh-market fruit types (68 round fruit types, 4 plum types, and 1 grape type).

Materials and Methods

PLANT MATERIAL. The tomato materials used in this study are elite modern tomato germplasm that were registered and stored by the UF/IFAS tomato breeding program (UF/IFAS, 2021) or

the NCSU tomato breeding program (NCSU, 2021); hereafter, UF tomato germplasm are termed Fla. and NCSU germplasm are termed NC. Both breeding programs are located in major fresh-market tomato-producing areas of the United States (USDA, 2016). Many germplasm used in this study have been commercialized (e.g., Gardner, 2006; Scott et al., 2008) and transferred to public/commercial tomato research programs (Florida Foundation Seed Producers, 2021; NCSU, 2021). All tomato materials were determinate [*sp/sp*, homozygous at the self-pruning locus (Barton et al., 1955)], fresh-market inbred germplasm. Two different segregating populations used to construct individual genetic maps were developed from a cross between the two elite modern *S. lycopersicum* tomatoes, Fla. 8653 and Fla. 8916 [hereafter referred to as Fla. population (Bhandari and Lee, 2021)] and from a cross between the two elite modern *S. lycopersicum* tomatoes, NC 13 and NC 3 (hereafter referred to as the NC population). A total of 15 [5 Fla. (4 plum fruit type, and 1 grape) and 10 NC (all round fruit type)] core fresh-market inbred tomatoes that were nominated by either the UF/IFAS tomato breeding program or NCSU and available as of Mar. 2019 were obtained (Supplemental Table 1).

WHOLE-GENOME SEQUENCING. Plants were grown as previously described [“Phenotype analysis” section in Lee et al. (2018)]. A single plant of each tomato was used for DNA extraction. Genomic DNA was extracted with a DNeasy Plant Mini Kit (Qiagen, Germantown, MD). We performed WGS of 110 plants (95 F₂ plants randomly selected from a segregating progeny in the NC population and 15 core tomatoes) using an Illumina HiSeq instrument (Illumina, San Diego, CA) as described in our previous study [“Whole-genome sequencing” section in Bhandari and Lee (2021)]. The sequencing libraries were sequenced on average 6 Gb reads for F₂ plant and 23 Gb for core inbred tomatoes, as described in our previous study [“Whole-genome sequencing” section in Bhandari and Lee (2021)]. In addition, Illumina whole-genome raw read datasets obtained from a previous study (Bhandari and Lee, 2021) were as follows: 93 F₂ plants randomly selected from a segregating progeny in the Fla. population and 58 core fresh-market tomatoes.

DETECTING SNPs. SNP detection was performed as described in a previous study [“Detecting single nucleotide polymorphisms” section in Bhandari and Lee (2021)]. Once SNP detection was completed for all plants, three separate SNP datasets were prepared for analysis. First, after removing the SNP calls supported by fewer than six reads in the WGS data from the F₂ plants and four parental tomatoes, SNP sites with heterozygosity in any parental tomatoes, monomorphic alleles, or more than 60% missing genotype codes due to variability in coverage were removed before analysis. To calculate the SNP and allelic concordance between F₂ and parental tomatoes (i.e., two parents used to create a biparental mating for the corresponding F₂

Table 1. Brief summary of three different tomato consensus maps.

Reference	Species	Population	Genotype source ^c
Doganlar et al., 2002	<i>Solanum lycopersicum</i> , <i>Solanum pimpinellifolium</i>	BC ₁ , BC ₂ , IBL ^y	RFLP
Robbins et al., 2011	<i>S. lycopersicum</i> , <i>S. pimpinellifolium</i>	F ₂	indels, RFLP, SNP, SSRs
This study	<i>S. lycopersicum</i>	F ₂	SNP

^zindels = insertions and deletions; RFLP = restriction fragment length polymorphism; SNP = single nucleotide polymorphism; SSRs = simple sequence repeats.

^yIBL = inbred backcross line.

population development), F_2 SNP alleles that did not show agreement with parental WGS calls were excluded. To avoid biased recombination information from imperfect allele segregation and/or missing genotypes, we selected the first common SNP site found in both populations (i.e., the SNP site agreement between two populations) per 10-kilo-base pair (kbp) interval across the genome, which yielded 335 informative polymorphic SNP sites (segregating SNP alleles; hereafter referred to as population data). Furthermore, we selected the first SNP site found in each population per 10-kbp interval across the genome in the case where the common SNP site was not available at any given 10-kbp interval, which yielded 6553 SNP sites (hereafter referred to as extended population data). Last, after removing SNP calls supported by fewer than six reads on the WGS data of 73 core fresh-market tomatoes, SNP calls across the tomatoes were grouped to create a SNP dataset (hereafter, inbred data).

CONSTRUCTION OF THE FRESH-MARKET TOMATO CONSENSUS GENETIC MAP. An individual genetic map was constructed as described in a previous study [“Construction of the large-fruited fresh-market tomato genetic map” section in Bhandari and Lee (2021)]. We estimated the degree of collinearity between the genetic and physical orders of markers in each genetic map as described in a previous study [“Construction of the large-fruited fresh-market tomato genetic map” section in Bhandari and Lee (2021)]. We constructed a fresh-market tomato consensus genetic map of 335 SNP markers using LPmerge (Endelman and Plomion, 2014). The maximum interval size K was chosen according to the tutorial referenced in Endelman and Plomion (2014). We then constructed an extended consensus genetic map of 6888 sites that combines both the population data and extended population data.

LINKAGE PANEL DEVELOPMENT. To create a fresh-market tomato linkage panel, the inbred data were mapped to the 335 SNP sites of the population data. Minor allele frequency [MAF (the least frequent allele at a given SNP site)] was also calculated.

The allelic agreement (e.g., chromosome 1, position 68611239, nucleotide C) between SNP calls in this study and the SolCAP array SNP calls [Table S2 in Sim et al. (2012a)] was confirmed as described in a previous study [“Linkage panel development” section in Bhandari and Lee (2021)].

GENETIC MAPPING OF THE TOMATO BRACHYTIC TRAIT USING THE LINKAGE PANEL. Two different mapping methods, linkage and association mapping, were applied to categorical phenotypic data for the tomato *brachytic* (*br*) trait (Barton et al., 1955; MacArthur, 1931). Two separate genotypic datasets were used in linkage mapping of the *br* trait in the Fla. population [Fla. 8916 carries *br* (*br/br*), Fla. 8653 does not carry the *br* allele]; 335 SNPs in the linkage panel (hereafter, 335 SNPs linkage mapping) and 3843 SNPs that combine both the 335 SNPs in the linkage panel and 3508 SNP sites found in the Fla. population (a subset of the extended population data) (hereafter, 3843 SNPs linkage mapping). We used the R environment (4.0.5) to implement the R/qtl package [version 1.48-1 (Arends et al., 2010)]; the single-QTL model using the function `cim()` was used to estimate the interval(s) for QTL.

Next, the Multiple Locus Mixed Linear Model (MLMM) within the Genome Association and Prediction Integrated Tool (GAPIT) version 3 (Wang and Zhang, 2020) was used to conduct the association analysis on 73 tomatoes in the linkage panel (hereafter, linkage panel association mapping). The allelic information in the linkage panel was used; the significant P value

threshold was 0.000149, which was 0.05% after Bonferroni multiple test correction. To validate the linkage panel association mapping, we conducted genome-wide association (GWA) analysis using a separate SNP dataset based on the inbred data (hereafter, GWA mapping). Of these 778,452,000 SNPs in the inbred data, a set of 265,910 SNPs with a missing genotype rate <5% was used for the GWA mapping. The Balding-Nichols kinship matrix and 10 principal components were included in the MLMM model to control population structure and individual relatedness. The significant P value threshold was 0.000000188, which was 0.05% after Bonferroni multiple test correction.

Tomatoes used in mapping were grown as previously described [“Phenotype analysis” section in Lee et al. (2018)]. The *br* trait was phenotyped as described in Lee et al. (2018). The phenotypic data used in this study can be found in Supplemental Table 2.

Results and Discussion

CONSTRUCTION OF THE FRESH-MARKET TOMATO CONSENSUS GENETIC MAP. To construct the consensus genetic map of fresh-market tomato that serves as a practical approximation to the distribution of recombination events in the tomato germplasm in breeding programs, we used two biparental F_2 segregating populations of fresh-market tomatoes as plant materials, which were nominated by public tomato breeding programs in the United States. After the initial SNP calling steps (“Detecting SNPs” section in Materials and Methods) and the comparison of parental line SNP calls and the corresponding F_2 population, a total of 75,806 SNP sites for the Fla. population and 1,233,164 for the NC population were identified. After selection of the one common SNP site per 10-kbp window, 5839 SNP sites for the Fla. population and 18,156 for the NC population were processed to correct errors in the genotypic datasets as described in a previous study (Hwang and Lee, 2019).

A complicating factor for the integration of two genetic maps in our study was the fact that not all SNP positions are physically evenly distributed over chromosomes. Consequently, a large proportion of the SNPs described previously then became unique positions in either population. More markers are still needed to obtain complete genome coverage, and thus, the integration of diverse genetic populations is the first necessary step toward achieving a high-resolution consensus genetic map while designing optimal marker selection criteria with respect to identifying those regions that are particularly devoid of common SNP sites in this study. Following genotypic dataset preprocessing and error corrections, a final set of 335 common SNPs was used to calculate the genetic distance between SNPs in each population. The total genetic map length of the Fla. population and NC population was estimated to be 943.2 and 947.1 cM, respectively, with chromosome 2 containing more than 40% of the total SNP markers and chromosomes 4, 5, 7, and 10 together representing less than 10% of the SNPs (Supplemental Table 3). It is interesting to note that an overall higher recombination rate was found along chromosomes 1, 2, and 3 of the NC population than those of the Fla. population (Fig. 1), although the overall recombination rate in the NC population was calculated at a similar level to that of the Fla. population (Supplemental Table 3). It has been suggested that there are differences in the overall chromosome-wide recombination rate along each chromosome between different modern tomato market classes [i.e., fresh-market class vs.

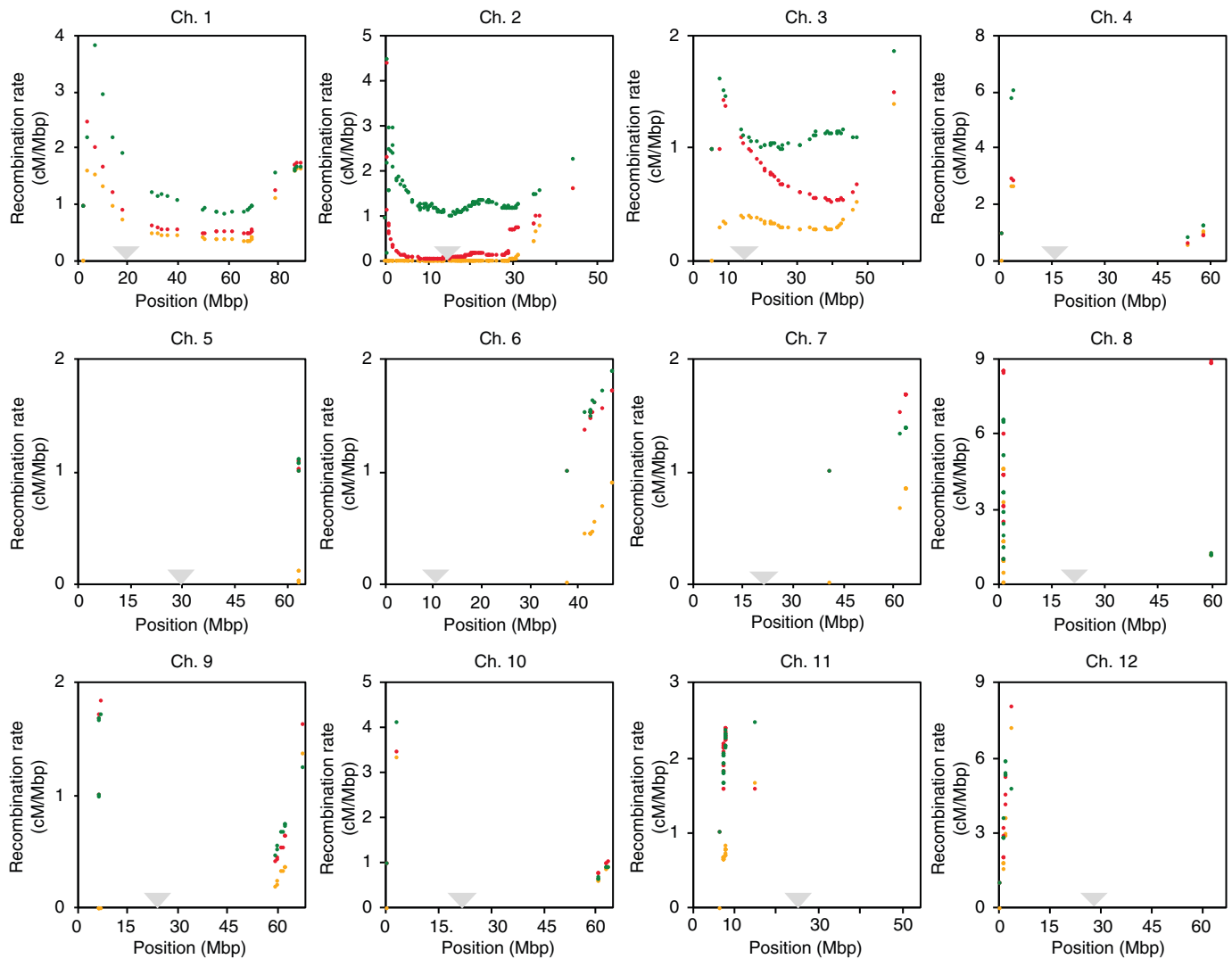


Fig. 1. Recombination rates of fresh-market tomato chromosomes (Ch.). Consensus genetic map, genetic map for Fla. 8653 \times Fla. 8916 (Fla. population), and genetic map for NC 13 \times NC 3 (NC population) are shown in orange, red, and green, respectively. The approximate centromere locations are depicted by gray triangles.

processing (Sim et al., 2012b)]. Although strong recombination suppression was reported in the pericentromeric regions within each tomato chromosome (Tomato Genome Consortium, 2012) and in the chromosomal region where interspecies crossings have occurred (Bhandari and Lee, 2021; de Haas et al., 2017; Sim et al., 2012a), we are unaware of chromosome-wide introgression of wild tomato into chromosomes 1, 2, and 3 of parental tomatoes in the Fla. population. Thus, our data indicate that the recombination rate may be variable between breeding populations for the same market class. It is unclear whether such a high overall chromosome-wide recombination rate in one of the breeding populations can be observed across all 12 tomato chromosomes, given that the recombination rates within a large proportion of the chromosome length, except for those three chromosomes, remained unresolved.

The consensus genetic map of 335 common SNP loci positioned on 12 chromosomes was estimated to be 737.3 cM (Table 2). The total genome coverage of the map was 466.9 Mbp, which represented 60% of the physical size of the reference tomato genome

assembly (Fernandez-Pozo et al., 2015). Previous studies reported that the total length of tomato genetic maps ranged from 867 to 2023.6 cM (Ashrafi et al., 2009; Bhandari and Lee, 2021; Chetelat et al., 2000; de Vicente and Tanksley, 1993; Gonda et al., 2019; Robbins et al., 2011; Sim et al., 2012a; Tanksley et al., 1992). However, genetic recombination in fresh-market tomato was not fully captured by the consensus map in this study, which is evident from the current genome coverage of the consensus map. The average resolution was 3.9 cM, and the average physical distance between flanking markers was 5.2 Mbp. There were a total of 19 genetic intervals that were larger than 10 cM, with the largest interval being 70.4 cM on chromosome 9. In a comparison between the genetic and physical orders of markers, no potential noncollinear markers were identified. In general, genetic and physical maps are collinear between close relatives in a species, but the degree of collinearity between them at the macrolevel (chromosome) and/or microlevel (gene) varies. Consensus marker orders between close species (e.g., modern and wild tomatoes) provide a valuable framework for utilization of such marker information from reference species to less

Table 2. Consensus genetic map of fresh-market tomato.

Chromosome no.	Consensus map										Extended consensus map ^z		
	Physical size (Mbp) ^y	Markers (no.)	Genetic size (cM)	Recombination rate (cM/Mbp)	Mean distance between flanking markers (cM)	Maximum distance between flanking markers (cM)	Genetic intervals (no.) ^x	Genome coverage (Mbp) ^w	Mean distance between flanking markers (Mbp)	Markers (no.)	Genetic size (cM)	Mean distance between flanking markers (cM)	Genome coverage (Mbp)
1	90.9	31	145.1	1.60	4.8	59.6	2	86.4 (0.95)	2.9	1,905	278.3	0.15	89.9
2	53.5	140	71.8	1.34	0.5	43.1	1	44 (0.82)	0.4	1,344	161.9	0.12	49.2
3	65.3	40	80.4	1.23	2.1	55.5	1	51.9 (0.79)	1.6	357	109.6	0.31	55.2
4	64.5	6	60.8	0.94	12.2	28.9	3	57.1 (0.86)	10.8	251	103.3	0.41	60.6
5	65.3	5	6.8	0.10	1.7	4.7	0	0.4 (0.01)	13.1	65	174.2	2.68	62.1
6	47.3	15	42.5	0.90	3.0	18.6	2	9.2 (0.20)	3.2	278	192.5	0.69	47.1
7	67.9	10	53.5	0.79	5.9	41.2	2	22.8 (0.34)	6.8	276	132.3	0.48	66.7
8	64.0	26	71.6	1.12	2.9	60.5	1	58.8 (0.92)	2.5	289	215.9	0.75	63.6
9	68.5	14	93.3	1.36	1.2	70.4	2	61.6 (0.90)	4.9	1,086	161.9	0.15	68.2
10	64.8	8	56.6	0.87	8.1	26.8	3	62.7 (0.97)	8.1	445	153.2	0.34	64.2
11	54.4	29	24.8	0.46	0.9	18.3	1	8.2 (0.15)	1.9	306	155.9	0.51	46.9
12	66.7	11	30.1	0.45	3.0	22.7	1	3.8 (0.06)	6.1	286	158.9	0.56	61.7
<i>Average</i>				0.93	3.9	37.5			5.2			0.60	
<i>Total</i>	773.1	335	737.3				19	466.9 (0.60)		6,888	1,997.9		735.4

^zThe map constructed after integration of 6553 markers that have been detected on at least one of two populations with the common 335 markers.

^yThe tomato genome version SL4.0 (Fernandez-Pozo et al., 2015) was used as the reference.

^xDefined by larger than 10-cM distance between flanking markers.

^wDefined by a physical distance (Mbp) between the first marker to the last one. Numbers in parentheses represent the proportion of the genome coverage to the corresponding chromosome size (Mbp).

well-understood genetics/breeding materials. The current tomato genome assembly was built on the Heinz 1706 processing tomato cultivar (Tomato Genome Consortium, 2012), which has a different market class than the tomato materials used in this study. Several additional draft tomato genome assemblies from long-read data were released recently, including at least one fresh-market tomato (Alonge et al., 2020) and wild tomatoes (Mazo-Molina et al., 2019; Schmidt et al., 2017; Wang et al., 2020). Furthermore, the genomes of different modern tomato market classes are currently being assembled (NRGene, 2021). Thus, our consensus map can be used as a framework as described previously for tomato genetics and breeding while taking better advantage of such new genomic resources.

CONSTRUCTION OF THE EXTENDED FRESH-MARKET TOMATO CONSENSUS GENETIC MAP. To integrate the 6553 SNP sites that were detected in either of two populations with the consensus map, new individual genetic maps were constructed. For the Fla. population, a final set of 3843 SNP sites that combines both the 335 common SNP sites and 3508 unique sites found in the population were processed for map construction. Of 3843 SNP sites, 2404 sites were in common with the SNP marker sites identified for the previous fresh-market tomato genetic map (Bhandari and Lee, 2021). The total map length was estimated to be 2413.2 cM (Supplemental Tables 4 and 5), which was 20% larger than expected based on previous estimates of genetic map length for this population (Bhandari and Lee, 2021). Differences between the current and previous maps were further assessed by comparing cumulative genetic maps, revealing that several narrow regions of high recombination rates were apparent in the telomere-proximal regions (e.g., chromosomes 6, 8, 11, and 12) in the current map, and the overall recombination rate was increased along a few chromosomes (e.g., chromosomes 2 and 9) (Supplemental Fig. 1). The limited number of plant samples in our study, misplaced or incorrectly ordered markers due to the complex nature of the genomes between tomatoes, and/or potential incomplete SNP calls in the population data, together with the concentration of SNPs in short regions [note there was a 14% SNP increment compared with the proportion of SNPs residing on chromosomes 2 and 9 in the previous study (Bhandari and Lee, 2021)], may have affected the map length, as discussed in our previous study (Bhandari and Lee, 2021). The two chromosomes 2 and 9 together contained more than 55% of the total SNP markers, and chromosomes 5, 6, 11, and 12 together represented less than 10% of the SNPs, revealing overall similar patterns for the number of markers as those reported in a previous study (Bhandari and Lee, 2021). A total of 101 potential noncollinear markers were identified. For the NC population, a final set of 3380 SNP sites with 3045 unique sites found in the population estimated the map length to be 2147.7 cM (Supplemental Tables 5 and 6). Unlike the Fla. population, chromosome 1 of the NC population showed the highest number of polymorphic sites, which accounted for 49.4% of the total SNP sites. A total of 52 potential noncollinear markers were identified in this population.

The resulting extended fresh-market tomato consensus map of 6888 SNPs was estimated to be 1997.9 cM (Table 2), which was compatible with the 2023.6 cM of the large-fruited fresh-market tomato genetic map (Bhandari and Lee, 2021). The average resolution was reduced by 85% relative to the consensus map of 335 markers. The total genome coverage of the extended consensus map was 735.4 Mbp, which represented more than 95% of the physical size of the reference tomato genome assembly.

DEVELOPMENT OF A FRESH-MARKET TOMATO LINKAGE PANEL OF 335 SNP LOCI. The inbred data were projected onto the SNP sites obtained from the consensus genetic map, and the Universal Fresh-Market Tomato Linkage Panel (UFTLP 2.0), which contained 335 SNP loci, was created (Supplemental Table 1). We detected 173 UFTLP 2.0 loci in UFTLP 1.0 (Bhandari and Lee, 2021). Among the 335 loci, 101 were multiallelic, and the rest were biallelic. There was an uneven distribution of MAFs in the panel, with a minimum proportion of 0 and a maximum of 0.47 (Table 3). The proportion of SNPs with a MAF from 0.4 to 0.5 was more than two times larger than that of the second largest MAF class from 0.3 to 0.4. The mean MAF was 0.35.

To determine how many of the 335 loci are expected to be polymorphic in a randomly selected biparental population, a complete pairwise comparison was applied to the 73 core fresh-market tomatoes. A total of 148 ± 2 loci (mean \pm 95% confidence interval) [with core tomato2 and core tomato26 showing the highest (166.0) and lowest (130.0) loci, respectively] were expected to be polymorphic in populations created by mating any two tomatoes (Supplemental Table 1). The mean 44% polymorphism rate among core fresh-market germplasm provided intriguing evidence of DNA sequence diversity in such tomatoes, especially in applied tomato genetics and U.S. fresh-market tomato breeding practices, given that many tomatoes used in this study have been commercialized (e.g., Gardner, 2006; Scott et al., 2008) and transferred to public/commercial tomato research programs (Florida Foundation Seed Producers, 2021; NCSU, 2021). As a resource for the scientific community, we have provided the full allelic information of 73 core tomatoes in UFTLP 2.0 (Supplemental Table 1). We also assessed the allelic agreement between the UFTLP 2.0 SNP calls and SolCAP SNP calls (note that the SolCAP array carries SNP calls for 31 tomatoes of the 73 UFTLP 2.0 tomatoes, which excluded two derivatives of originally released germplasm). We detected three UFTLP 2.0 SNP loci in the SolCAP array (Supplemental Table 1), with 81 UFTLP 2.0 alleles that agreed with SolCAP SNP calls and two alleles that did not. In addition, one allele remained unresolved due to a heterozygous SolCAP SNP call.

GENETIC MAPPING OF THE TOMATO BRACHYTIC TRAIT USING THE LINKAGE PANEL. SNP data in linkage panels [the linkage panel data in this study as well as the linkage panel version 1.0 in Bhandari and Lee (2021)] can be used to genotype diverse tomato samples. To demonstrate that the linkage panel can be useful to identify genetic associations with traits in fresh-market tomatoes, we performed two different genetic mapping analyses using the linkage panel. The 335 SNPs linkage mapping identified an SNP signal [SNP UFTLP2_19 (Supplemental Table 1)]

Table 3. Distribution of the minor allele frequencies in fresh-market tomato linkage panel 2.0.

Minor allele frequency class	Markers (no.)	Proportion
0–0.1	0	0.00
0.1–0.2	52	0.16
0.2–0.3	59	0.18
0.3–0.4	66	0.20
0.4–0.5	158	0.47
Total	335	1.00

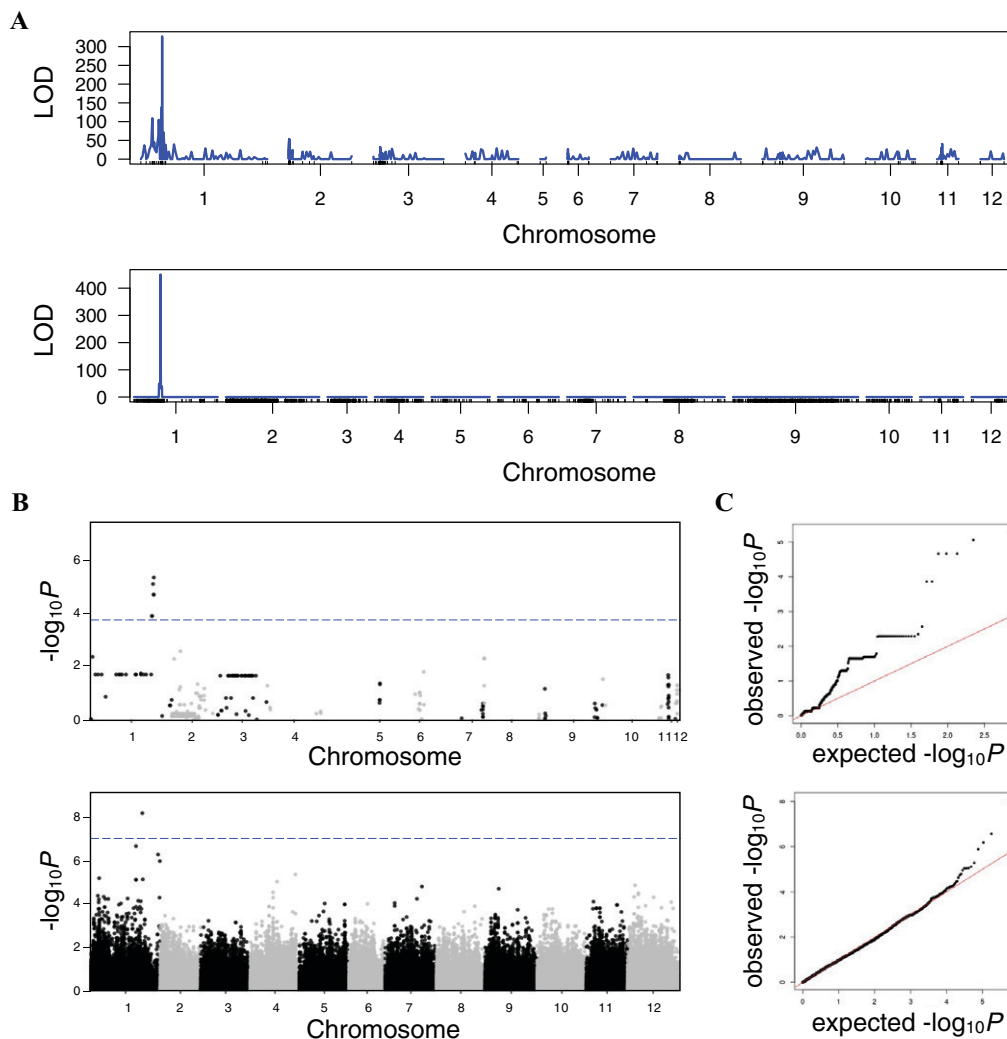


Fig. 2. Genetic mapping of the tomato *brachytic* trait. (A) Linkage mapping: logarithm of the odds (LOD) ratio plots of single nucleotide polymorphisms (SNPs) associated with the *brachytic* trait in the F_2 segregating population [top plot, 335 SNPs in the linkage panel; bottom plot, 3843 SNPs that combines both the 335 SNPs in the linkage panel and 3508 SNP sites found in the F_2 population from Fla. 8653 \times Fla. 8916 (Fla. population)]. The x-axis shows chromosomal position. Black tick markers on the x-axis indicate genetic locations of SNP markers. (B) Association mapping: Manhattan plots of SNPs associated with the *brachytic* trait in 73 core fresh-market inbred tomatoes from the linkage panel (top plot, 335 SNPs in the linkage panel; bottom plot, genome-wide 265,910 SNPs). The y-axis indicates $-\log_{10} P$ values for association signals. The dashed horizontal line indicates the calculated threshold $-\log_{10} P$ value (3.83 and 6.73; top and bottom plots, respectively) for determining a significant association. (C) Quantile-quantile plots for association mapping (top plot, 335 SNPs in the linkage panel; bottom plot, genome-wide 265,910 SNPs).

with a significant effect (logarithm of the odds ratio for linkage >300) on the *brachytic* trait (top plot in Fig. 2A). The signal has a bp position located ≈ 1.3 -Mbp away from the fine-mapped *br* interval (Lee et al., 2018). Interestingly, a strongly significant 3843 SNPs linkage mapping signal was detected at 68,347,169 bp on chromosome 1, which fell within the fine-mapped *br* interval spanned a 763.6-kbp region (i.e., 67,891,398–68,654,954 bp on chromosome 1 on the SL4.0 version of the tomato genome assembly) (bottom plot in Fig. 2A). The SNP signal from the 3843 SNPs linkage mapping, whose position corresponded to the respective fine-mapped *br* interval, could be attributable to an increased number of SNPs in the 3843 SNPs linkage mapping.

Although the current linkage panel is subjected to improvement through introgressions from diverse populations/germplasm and marker densities, the linkage panel identified association signals (i.e., SNPs that passed the test threshold) for

brachytic from the linkage panel association mapping (top plot in Fig. 2B). A single 835.0-kbp region comprising five association-detected SNPs (68,611,239–69,446,203 bp on chromosome 1) overlapped the telomere-proximal region of the fine-mapped *br* interval. The GWA mapping, however, exhibited the strongest SNP signal 417.2-kbp away from the fine-mapped *br* (bottom plot in Fig. 2B). We also noted two SNPs (with P values of 0.000000272 and 0.00000065) flanking the fine-mapped *br* interval in the 29.0-Mbp region, but both signals did not exceed a specified threshold (0.000000188) for the Bonferroni correction in this study. In the linkage panel association mapping, the sharp deviation above an expected $-\log_{10} P$ value of ≈ 0.25 was observed (top plot in Fig. 2C); however, correction for these by genome-wide SNP markers in the GWA mapping reduced such deviations to the identity line (bottom plot in Fig. 2C).

Conclusions

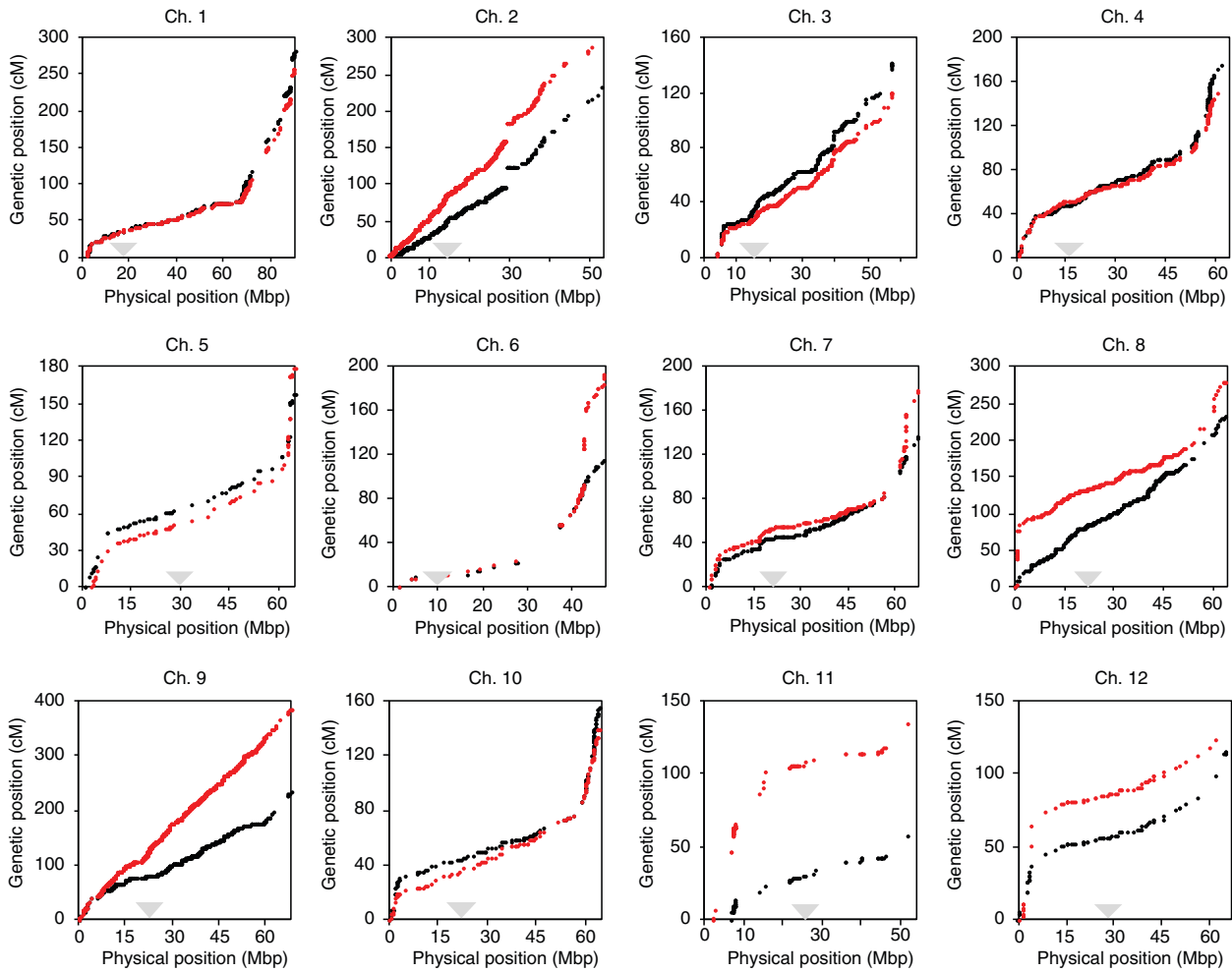
Tomato is an important crop, and this is the first study that reports the fresh-market tomato consensus genetic map of 335 SNP markers. A high degree of collinearity between the genetic position of the SNP markers on the consensus map and their physical positions on the tomato reference genome, together with the second version of a fresh-market tomato linkage panel UFTLP 2.0, will contribute to coordinating the implementation of the SNP markers (such as genotyping platforms and genomic association map) for tomato genetics and breeding.

Recently developed next-generation sequencing technologies provide massive DNA sequence information at lower cost. However, such technologies are still relatively expensive. Thus, mass de novo whole-genome resequencing and reduced representation sequencing of plant species will require still more advanced and lower cost technologies than those currently on the market (Hudson, 2008). A consensus genetic map that depicts the position (both marker order and distance between flanking markers) of genetic markers for various plant materials from a single species can provide practical, genetic data resources to researchers. Unlike the physical map or even complete genome sequence, the genetic map provides genetic recombination, which is an essential component of current crop improvement efforts. The growing availability of whole-genome information, together with developments in (consensus) genetic maps, could facilitate crop improvement.

Literature Cited

- Alonge, M., X. Wang, M. Benoit, S. Soyk, L. Pereira, L. Zhang, H. Suresh, S. Ramakrishnan, F. Maumus, D. Ciren, Y. Levy, T.H. Harel, G. Shalev-Schlosser, Z. Amsellem, H. Razifard, A.L. Caicedo, D.M. Tieman, H. Klee, M. Kirsche, S. Aganezov, T.R. Ranallo-Benavidez, Z.H. Lemmon, J. Kim, G. Robitaille, M. Kramer, S. Goodwin, W.R. McCombie, S. Hutton, J. Van Eck, J. Gillis, Y. Eshed, F.J. Sedlazeck, E. van der Knaap, M.C. Schatz, and Z.B. Lippman. 2020. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell* 182:145–161, <https://doi.org/10.1016/j.cell.2020.05.021>.
- Arends, D., P. Prins, R.C. Jansen, and K.W. Broman. 2010. R/qtl: High-throughput multiple QTL mapping. *Bioinformatics* 26:2990–2992, <https://doi.org/10.1093/bioinformatics/btq565>.
- Ashrafi, H., M. Kinkade, and M.R. Foolad. 2009. A new genetic linkage map of tomato based on a *Solanum lycopersicum* x *S. pimpinellifolium* RIL population displaying locations of candidate pathogen response genes. *Genome* 52:935–956, [0.1139/g09-065](https://doi.org/10.1139/g09-065).
- Barton, D.W., L. Butler, J.A. Jenkins, C.M. Rick, and P.A. Young. 1955. Rules for nomenclature in tomato genetics (includes a list of known genes). *J. Hered.* 46:22–76, <https://doi.org/10.1093/oxfordjournals.jhered.a106504>.
- Bhandari, P. and T.G. Lee. 2021. A genetic map and linkage panel for the large-fruited fresh-market tomato. *J. Amer. Soc. Hort. Sci.* 146:125–131, <https://doi.org/10.21273/JASHS04999-20>.
- Blanca, J., J. Montero-Pau, C. Sauvage, G. Bauchet, E. Illa, M.J. Díez, D. Francis, M. Causse, E. van der Knaap, and J. Cañizares. 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics* 16:257, <https://doi.org/10.1186/s12864-015-1444-1>.
- Chetelat, R.T., V. Meglic, and P. Cisneros. 2000. A genetic map of tomato based on BC₁ *Lycopersicon esculentum* x *Solanum lycopersicoides* reveals overall synteny but suppressed recombination between these homeologous genomes. *Genetics* 154:857–867.
- de Haas, L.S., R. Koopmans, C. Lelivelt, R. Ursem, R. Dirks, and G. Velikkakam James. 2017. Low-coverage resequencing detects meiotic recombination pattern and features in tomato RILs. *DNA Res.* 24:549–558, <https://doi.org/10.1093/dnares/dsx024>.
- de Vicente, M.C. and S.D. Tanksley. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134:585–596.
- Doganlar, S., A. Frary, H.M. Ku, and S.D. Tanksley. 2002. Mapping quantitative trait loci in inbred backcross lines of *Lycopersicon pimpinellifolium* (LA1589). *Genome* 45:1189–1202, <https://doi.org/10.1139/g02-091>.
- Endelman, J.B. and C. Plomion. 2014. LPmerge: An R package for merging genetic maps by linear programming. *Bioinformatics* 30:1623–1624, <https://doi.org/10.1093/bioinformatics/btu091>.
- Fernandez-Pozo, N., N. Menda, J.D. Edwards, S. Saha, I.Y. Tecle, S.R. Strickler, A. Bombarely, T. Fisher-York, A. Pujar, H. Foerster, A. Yan, and L.A. Mueller. 2015. The Sol Genomics Network (SGN) - From genotype to phenotype to breeding. *Nucleic Acids Res.* 43:D1036–D1041, <https://doi.org/10.1093/nar/gku1195>.
- Florida Foundation Seed Producers. 2021. Florida Foundation Seed Producers, Inc. 16 Sept. 2021. <www.ffsp.net>.
- Food and Agriculture Organization of the United Nations. 2016. Value of agricultural production. 16 Sept. 2021. <www.fao.org/faostat/en/#data/QC/metadata>.
- Gardner, R.G. 2006. ‘Plum Crimson’ fresh-market plum tomato hybrid and its parents, NC EBR-7 and NC EBR-8. *HortScience* 41:259–260.
- Gonda, I., H. Ashrafi, D.A. Lyon, S.R. Strickler, A.M. Hulse-Kemp, Q. Ma, H. Sun, K. Stoffel, A.F. Powell, S. Futrell, T.W. Thannhauser, Z. Fei, A.E. Van Deynze, L.A. Mueller, J.J. Giovannoni, and M.R. Foolad. 2019. Sequencing-based bin map construction of a tomato mapping population, facilitating high-resolution quantitative trait loci detection. *Plant Genome* 12:10.3835, <https://doi.org/10.3835/plantgenome2018.02.0010>.
- Hudson, M.E. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Resour.* 8:3–17, <https://doi.org/10.1111/j.1471-8286.2007.02019.x>.
- Hwang, S. and T.G. Lee. 2019. Correcting pervasive errors in genotypic datasets to develop genetic maps. *Agronomy (Basel)* 9:196, <https://doi.org/10.3390/agronomy9040196>.
- Lee, T.G., S.F. Hutton, and R. Shekasteband. 2018. Fine mapping of the *brachytic* locus on the tomato genome. *J. Amer. Soc. Hort. Sci.* 143:239–247, <https://doi.org/10.21273/JASHS04423-18>.
- Lukyanenko, A.N. 1991. Disease resistance in tomato, p. 99–120. In: G. Kallou (ed.). *Genetic improvement of tomato*. Springer-Verlag, Berlin, Germany.
- MacArthur, J.W. 1931. Linkage studies with the tomato. III Fifteen factors in six groups. *Trans. Royal Can. Inst.* 18:1–19.
- Mazo-Molina, C., S. Mainiero, S.R. Hind, C.M. Kraus, M. Vachev, F. Maviane-Macia, M. Lindeberg, S. Saha, S.R. Strickler, A. Feder, J.J. Giovannoni, C.D. Smart, N. Peeters, and G.B. Martin. 2019. The *Ptr1* locus of *Solanum lycopersicoides* confers resistance to race 1 strains of *Pseudomonas syringae* pv. *tomato* and to *Ralstonia pseudosolanacearum* by recognizing the type III effectors AvrRpt2 and RipBN. *Mol. Plant Microbe Interact.* 32:949–960, <https://doi.org/10.1094/MPMI-01-19-0018-R>.
- North Carolina State University. 2021. NCSU tomato breeding program. 16 Sept. 2021. <<https://mountainhort.ces.ncsu.edu/fresh-market-tomato-breeding/>>.
- NRGene. 2021. Tomato. 16 Sept. 2021. <www.nrgene.com/solutions/consortia/tomato>.
- Robbins, M.D., S.C. Sim, W. Yang, A. Van Deynze, E. van der Knaap, T. Joobeur, and D.M. Francis. 2011. Mapping and linkage disequilibrium analysis with a genome-wide collection of SNPs that detect polymorphism in cultivated tomato. *J. Expt. Bot.* 62:1831–1845, <https://doi.org/10.1093/jxb/erq367>.
- Rodríguez, G.R., S. Muños, C. Anderson, S.C. Sim, A. Michel, M. Causse, B.B. Gardener, D. Francis, and E. van der Knaap. 2011. Distribution of *SUN*, *OVATE*, *LC*, and *FAS* in the tomato germplasm

- and the relationship to fruit shape diversity. *Plant Physiol.* 156:275–285, <https://doi.org/10.1104/pp.110.167577>.
- Schmidt, M.H., A. Vogel, A.K. Denton, B. Istace, A. Wormit, H. van de Geest, M.E. Bolger, S. Alseekh, J. Maß, C. Pfaff, U. Schurr, R. Chetelat, F. Maumus, J.-M. Aury, S. Koren, A.R. Fernie, D. Zamir, A.M. Bolger, and B. Usadel. 2017. De novo assembly of a new *Solanum pennellii* accession using nanopore sequencing. *Plant Cell* 29:2336–2348, <https://doi.org/10.1105/tpc.17.00521>.
- Scott, J.W., E.A. Baldwin, H.J. Klee, J.K. Brecht, S.M. Olson, J.A. Bartz, and C.A. Sims. 2008. Fla. 8153 hybrid tomato; Fla. 8059 and Fla. 7907 breeding lines. *HortScience* 43:2228–2230, <https://doi.org/10.21273/HORTSCI.43.7.2228>.
- Scott, J.W., J.R. Myers, and P.S. Boches. 2013. Classical genetics and traditional breeding, p. 37–73. In: B.E. Liedl, J.A. Labate, J.R. Stommel, A. Slade, and C. Kole (eds.). *Genetics, genomics and breeding of tomato*. CRC Press, Boca Raton, FL.
- Sim, S.C., G. Durstewitz, J. Plieske, R. Wieseke, M.W. Ganal, A. Van Deynze, J.P. Hamilton, C.R. Buell, M. Causse, S. Wijeratne, and D.M. Francis. 2012a. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS One* 7:e40563, <https://doi.org/10.1371/journal.pone.0040563>.
- Sim, S.C., A. Van Deynze, K. Stoffel, D.S. Douches, D. Zarka, M.W. Ganal, R.T. Chetelat, S.F. Hutton, J.W. Scott, R.G. Gardner, D.R. Panthee, M. Mutschler, J.R. Myers, and D.M. Francis. 2012b. High-density SNP genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding. *PLoS One* 7:e45520, <https://doi.org/10.1371/journal.pone.0045520>.
- Tanksley, S.D. 2004. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell* 16:S181–S189, <https://doi.org/10.1105/tpc.018119>.
- Tanksley, S.D., M.W. Ganal, J.P. Prince, M.C. de Vicente, M.W. Bonierbale, P. Broun, T.M. Fulton, J.J. Giovannoni, S. Grandillo, and G.B. Martin. 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141–1160.
- Tanksley, S.D. and C.M. Rick. 1980. Isozymic gene linkage map of the tomato: Applications in genetics and breeding. *Theor. Appl. Genet.* 57:161–170, <https://doi.org/10.1007/BF00279708>.
- The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635–641, <https://doi.org/10.1038/nature11119>.
- University of Florida, Institute of Food and Agricultural Sciences. 2021. UF/IFAS tomato breeding program. 16 Sept. 2021. <<https://tombreeding.ifas.ufl.edu>>.
- U.S. Department of Agriculture. 2016. Tomatoes. 16 Sept. 2021. <www.ers.usda.gov/topics/crops/vegetables-pulses/tomatoes>.
- Wang, J., Z. Zhang. 2020. GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *bioRxiv* 2020.11.29.403170, doi: 10.1101/2020.11.29.403170.
- Wang, X., L. Gao, C. Jiao, S. Stravoravdis, P.S. Hosmani, S. Saha, J. Zhang, S. Mainiero, S.R. Strickler, C. Catala, G.B. Martin, L.A. Mueller, J. Vrebalov, J.J. Giovannoni, S. Wu, and Z. Fei. 2020. Genome of *Solanum pimpinellifolium* provides insights into structural variants during tomato breeding. *Nat. Commun.* 11:5817, <https://doi.org/10.1038/s41467-020-19682-0>.



Supplemental Fig. 1. Cumulative genetic maps of fresh-market tomato chromosomes (Ch.). The genetic maps of F_2 population from Fla. 8653 \times Fla. 8916 in this study and the previous study (Bhandari and Lee, 2021) are shown in red and black, respectively. The approximate centromere locations are depicted by gray triangles.