A Genetic Map and Linkage Panel for the Large-fruited Fresh-market Tomato

Prashant Bhandari

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

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 Molecular and Cellular Biology Graduate Program, University of Florida, Gainesville, FL 32611

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

ABSTRACT. Genetic maps saturated with genetic markers are useful for genetic research and crop breeding; however, the genetic map for the large-fruited fresh-market tomato (*Solanum lycopersicum*) has never been constructed, and the recombination frequency between DNA fragments is only partly understood for fresh-market tomato. We constructed a novel fresh-market tomato genetic map by using 3614 single nucleotide polymorphism (SNP) markers and a 93 F₂ segregating progeny derived from a cross between two United States large-fruited fresh-market tomato lines. The average distance between markers was less than 1 cM, and substantial recombination densities between markers were observed across the approximate centromere locations. A linkage panel for large-fruited fresh-market tomato was also established using the combined dataset of the genetic map and 58 SNP-genotyped core tomato lines. The allelic information in the linkage panel will be a significant resource for both tomato genetics and future breeding approaches.

Genetic improvements in crop plants have been achieved by mating selected parents and assessing the performance of the progeny. This method requires breeders or geneticists to assemble desirable genes. Genetic maps depict the location of genetic markers on chromosome(s), with the distance between two flanking markers on the same chromosome dependent on genetic recombination. In genetic maps, genes with known functions or DNA fragments can be used as genetic markers and their inheritance can be traced. Genetic maps saturated with genetic markers are useful for genetic research to identify new genes/quantitative trait loci (QTL), and for genomics research concerned with genome sequence. A complete genetic map that covers all chromosomes within a species genome is important because it provides a resource for developing a consensus genetic map that can be a common framework for mapping studies. Thus, accurate and high-resolution genetic maps are critical for crop improvement.

The most common genetic markers used in recent years have been DNA markers, but morphological markers are still used for a number of traits [e.g., tomato jointless pedicel (Butler, 1936)], which can be readily validated by observation. Recent advances in genotyping technologies have made it possible to create genetic maps and exploit genetic linkage between markers and traits with unprecedented resolution. For example, SNPs identified by high-throughput sequencing serve as markers of the associations between genotypes and phenotypes (Hudson, 2008).

Tomato (*Solanum lycopersicum*) is the most valuable horticultural crop worldwide (Food and Agriculture Organization of the United Nations, 2016) and provides micronutrients to the human diet (U.S. Department of Agriculture, 2016). Fresh-market and processing tomato are the two most commonly consumed types of tomatoes and account for more than \$2.6 billion in annual farm cash receipts in the United States alone, with more than \$1.0 billion of this value from fresh-market tomatoes (U.S. Department of Agriculture, 2016). Specifically, large-fruited fresh-market tomatoes cultivated in the United States (also called round tomato or beefsteak tomato) represent a unique type of tomato fruit class that is bred for direct consumption. These tomatoes account for more than \$600 million per year in farm cash receipts (Florida Tomato Exchange, 2020). In the United States, ≈40% of large-fruited fresh-market tomatoes are produced in Florida (up to 100% during the winter and early spring seasons) (Florida Tomato Exchange, 2020). Tomato yield has increased steadily throughout the years due to an increase in their genetic potential and advances in agronomic practices; however, a significant improvement in horticultural performance is still necessary, especially given the predicted shifts in climate, pathogen outbreaks, limited performance associated with soil fumigation practices, and new challenges from national/international producers (Guan et al., 2017). Therefore, it is necessary to understand the mechanisms of biotic/abiotic stress resistance and increase genetic diversity in tomato breeding to produce germplasm with a broader range of tolerance to different stressors and marketable traits.

There is no genetic map for the large-fruited fresh-market tomato, and genetics and breeding of the fresh-market tomato are currently based on maps originally constructed for the purpose of studying the inheritance and variation of biological traits rather than for fresh-market tomato research applications (Ashrafi et al., 2009; Chetelat et al., 2000; Gonda et al., 2019; Robbins et al., 2011; Sim et al., 2012a). Furthermore, the Solanaceae

Received for publication 22 Sept. 2020. Accepted for publication 6 Jan. 2021. Published online 5 February 2021.

This work is supported by the Florida Tomato Committee (T.G.L.) and USDA National Institute of Food and Agriculture (Hatch project FLA-GCC-005550; T.G.L.).

We thank Samuel F. Hutton for sharing tomato genetic 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/>).

Coordinated Agricultural Project (SolCAP) array (Sim et al., 2012a) and its derivatives [e.g., Tomato Genotyping Library (LGC Biosearch Technologies, Teddington, UK)] are currently used to genotype tomato, especially those bred in the United States, because it can be used as a readily accessible “entry point” for relatively inexpensive genotyping. However, the genetic variation in the large-fruited fresh-market tomato has not been well-captured by the array, which is evident from the low number of polymorphic sites and the presence of large genomic sequence regions without markers (data not shown). Therefore, this array is not readily suited for QTL linkage mapping and/or association studies, particularly genotyping populations derived from crosses between two large-fruited fresh-market tomatoes (S.F. Hutton, personal communication). The objectives of our project were to create a genetic map for large-fruited fresh-market tomato and a linkage panel for the core lines with two goals to improve our knowledge of the genetic recombination in these tomatoes, as well as provide loci applicable to independent linkage/association mappings for use in a linkage panel. To our knowledge, no prior biparental population between two large-fruited fresh-market tomatoes has been used to study genetic recombination. This work will be applicable to breeding and applied genetics programs.

Materials and Methods

PLANT MATERIAL. The term “line” is used here to mean tomato ($2n = 2x = 24$) genetic material that was registered and stored by the University of Florida (UF) Institute of Food and Agricultural Sciences (IFAS) tomato breeding program (UF/IFAS, 2020). All Florida tomato lines (hereafter, Florida tomato lines are referred to as “Fla.”) and populations were developed and maintained by the UF/IFAS tomato breeding program. A mapping population used to construct the genetic map was developed from a cross between the two elite *S. lycopersicum* lines, Fla. 8653 and Fla. 8916 (Fig. 1). Both lines were large-fruited, determinate [*sp/sp*, homozygous at the self-pruning locus (Barton et al., 1955)], fresh-market inbred lines. A total of 58 core fresh-market inbred lines that were nominated by the UF/IFAS tomato breeding program and available as of Feb. 2019 were obtained (Supplemental Table S1).

WHOLE-GENOME SEQUENCING. Plants were grown as previously described [Phenotype analysis section in the work by Lee

et al. (2018)]. Genomic DNA was extracted with a DNeasy Plant Mini Kit (Qiagen, Germantown, MD). A single plant of each tomato line was used for DNA extraction. We performed whole-genome sequencing (WGS) for 151 plants (93 F_2 plants randomly selected from a progeny derived from a cross between Fla. 8653 and Fla. 8916 and 58 core lines) using an Illumina HiSeq instrument (Illumina, San Diego, CA) with 150-bp paired-end sequencing at Novogene (Beijing, China). Sequencing libraries were prepared with a polymerase chain reaction-free NEBNext Ultra II DNA Library Prep Kit (NEB, Ipswich, MA) with an insert size of 350 bp. The libraries were sequenced on average 6 Gb (low-coverage sequence ranges from 5 to $9 \times$ depth per sequenced base) for F_2 plant reads, as well as 23 Gb (high-coverage sequence ranges from 22 to $32 \times$ depth) for inbred lines, including the two parental lines of the F_2 population. An average Phred quality score of 30 or higher was produced.

DETECTING SINGLE NUCLEOTIDE POLYMORPHISMS. To develop a genetic map and linkage panel, we identified SNPs in reads from each line aligned to the SL4.0 version of the tomato genome assembly (Fernandez-Pozo et al., 2015) using the BWA-MEM algorithm [version 0.7.17 (Li and Durbin, 2010)]. To increase the reliability of distinguishing homozygous from heterozygous genotypes (note that our lowest coverage F_2 dataset had an average depth per sequenced base of $6\times$), the SNP calls supported by fewer than six reads were removed before further analysis. Linux system commands, SAMtools [version 1.9 (Li et al., 2009)] and BCFtools [version 1.9 (Li et al., 2009)] connected by an in-house shell script were used to compute and compare the identified SNPs and generate datasets for creating the genetic map and linkage panel.

Two 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 F_2 plants and two parental lines, SNP sites with heterozygosity of any parental lines were also removed before analysis. To calculate the SNP and allelic concordance between F_2 and parental lines, WGS data were obtained to compare F_2 plants SNP calls and parental lines SNP calls (see Whole-genome sequencing section in this article). The F_2 SNP alleles that did not show agreement with parental WGS calls were excluded. In addition, we removed the SNP sites with monomorphic alleles or more than 50% missing genotype codes due to variability in coverage. To avoid biased recombination information from imperfect allele segregation and/or missing genotypes, we selected the first SNP site per 10-kbp interval across the genome, which yielded 3614 informative polymorphic SNP sites (segregating SNP alleles in the mapping population; hereafter referred to as population data). Second, after removing SNP calls supported by fewer than six reads on the WGS data of 58 inbred lines, SNP calls across the lines were grouped to create a second SNP dataset (hereafter referred to as inbred data). In contrast to the workflow used to generate population data, inbred data did not require additional filtration steps.

CONSTRUCTION OF THE LARGE-FRUITED FRESH-MARKET TOMATO GENETIC MAP. To construct a genetic map of large-fruited fresh-market tomato, we used the R environment (version 3.5.0; R Foundation, Vienna, Austria) with the R/qtl package [version 1.44-9 (Arends et al., 2010)]. When the population data were acquired, potential errors in the genotypic datasets were corrected as previously described (Hwang and Lee, 2019). The genetic map was constructed as described in

Fig. 1. Extra-large fruit of two U.S. fresh-market tomato lines, Fla. 8653 (left) and Fla. 8916 (right), according to the U.S. Department of Agriculture standard size designations for tomatoes (U.S. Department of Agriculture, 2005); scale bar = 1 cm.

our previous study (Hwang and Lee, 2019), with modifications in linkage grouping steps and resolving marker orders: a logarithm of odds threshold of 6 and recombination frequency (Liu, 1998; Ott, 1999; Terwilliger and Ott, 1994) of 0.35 were used and the positions of SNP markers from the SL4.0 version of the reference tomato genome assembly were used. We then used the R/qtl function `ripple()` in a sliding window of size 5 to estimate the degree of collinearity between the genetic (i.e., our population-based linkage test) and physical (i.e., SNP positions from the tomato reference genome) orders of markers. The marker positions (both marker order and map distance) were optimized with regard to crossover counts and the genetic length of each chromosome. The presence of noncollinear markers was validated using independent software, ASMap [version 1.0-4 (Taylor and Butler, 2018)], JoinMap [version 4.1 (Stam, 1993)], and THREaD Mapper [version 1.0 (Cheema et al., 2008)].

To compare the map of the full list of SNPs identified in the population data and the map constructed using the down-sampling of the SNPs, an iterative approach (Liu, 1998; Ott, 1999; Terwilliger and Ott, 1994) involving a total of 10 iterations run to create a new map was used. For each iteration, SNPs that segregated ≥ 5 cM of each other were randomly chosen; then, the genetic length of each chromosome was assessed.

LINKAGE PANEL DEVELOPMENT. To create a large-fruited fresh-market tomato linkage panel, the inbred data were mapped to the 3614 SNP sites of the population data. Minor allele frequency [MAF (the least frequent allele at a given SNP site)] was also calculated.

To confirm the allelic agreement (e.g., chromosome 1, position 2260216, nucleotide G) between SNP calls in this study and the SolCAP array SNP calls (Supplemental Table S2 in Sim et al., 2012b), the probe sequences from the array were aligned to the reference genome using the BLASTn program (version 2.10.0) and Linux system commands connected by an in-house shell script. All shell scripts in this study and the R function `est.map()` were run on the UF/Research Computing Linux server (University of Florida, 2020).

Results

CONSTRUCTION OF THE LARGE-FRUITED FRESH-MARKET TOMATO GENETIC MAP. To construct the genetic map of the large-fruited fresh-market tomato, genotypic datasets were collected from an F₂ population derived from a cross between two large-fruited fresh-market tomato lines (Fla. 8653 $\times$ Fla. 8916) using WGS. After the initial SNP calling steps (see the Single nucleotide polymorphism detection section) and comparisons of F₂ plant SNP calls and parental line SNP calls, a total of 75,185 SNP sites were identified. After selection of the 1 SNP site per 10-kbp window, 11,812 SNP sites were processed to correct errors in the genotypic datasets (see the Construction of the large-fruited fresh-market tomato genetic map section). Following genotypic dataset preprocessing and error corrections, a final set of 3614 SNPs was used to calculate the genetic distance between SNPs using the R/qtl package. The total map length was estimated to be 2023.6 cM, with chromosomes 2 and 9 containing more than 40% of the total SNP markers and chromosomes 5, 6, 11, and 12 together representing less than 10% of the SNPs (Table 1, Supplemental Fig. S1). The average resolution (i.e., average genetic distance between

flanking markers) was 0.86 cM, and the average physical distance between flanking markers was 369.3 kbp. There was a total of 25 genetic intervals that were larger than 10 cM, with the largest one being 41.6 cM on chromosome 1. In a comparison between the genetic and physical orders of markers, a total of 68 potential noncollinear markers that formed 22 segments (a segment represents a group of consecutive markers) were identified (Table 1, Supplemental Table S1, Supplemental Fig. S1).

DISTRIBUTION OF RECOMBINATION RATES IN THE LARGE-FRUITED FRESH-MARKET TOMATO GENOME. The average recombination rate was 2.62 cM/Mbp (Table 1). The highest recombination rate reached between the 0 and 10-Mbp regions in most chromosomes (Fig. 2). As the location neared the approximate centromere, the rate moved downward (with the exception of chromosome 2), whereas the overall rate of recombination was >1 cM/Mbp. The distribution of the recombination rate on chromosome 2 reached its peak within the last 10-Mbp region of the chromosome, whereas the recombination rates within 10% of the chromosome length of chromosomes 6 and 11 remained unresolved.

EFFECT OF MARKER DOWN-SAMPLING ON GENETIC MAP LENGTH. To assess the effect of SNP marker down-sampling on genetic length, we reconstructed the genetic map by selecting a randomly reduced number of markers (mean, 487 markers) that were separated by at least 5-cM intervals and recalculated the map length. This resampling analysis produced an estimated map length that was reduced by a mean of 31% (range, 12% to 127%) relative to the map that used the full marker set, with chromosomes 2 and 11 showing the highest and lowest inflations, respectively (Supplemental Table S2).

DEVELOPMENT OF A LARGE-FRUITED FRESH-MARKET TOMATO LINKAGE PANEL OF 3614 SNP LOCI. The inbred data were projected onto the SNP sites obtained from the genetic map, and a University of Florida's Fresh-Market Tomato Linkage Panel version 1.0 (UFTLP 1.0) of 3614 SNP loci was created (Supplemental Table S1). Among the 3614 loci, 20 were multiallelic loci and the others were biallelic. There was an uneven distribution of MAFs in the panel, with a minimum proportion of 0.08 and a maximum proportion of 0.40 (Table 2). The proportion of SNPs with a MAF class of 0.3 to 0.4 was more than 1.5-times larger than that of the second largest MAF class of 0 to 0.1. The mean MAF was 0.27.

To determine how many of the 3614 loci are expected to be polymorphic in a randomly selected biparental population, a complete pairwise comparison was applied to the 58 core lines. A total of 1362 ± 26 loci (mean $\pm$ 95% confidence interval) were expected to be polymorphic in populations created by mating any two lines. We also assessed the allelic agreement between the UFTLP 1.0 SNP calls and SolCAP SNP calls (note that the SolCAP array carries SNP calls for 28 lines of the 58 UFTLP 1.0 tomato lines, which excluded two derivatives of originally released lines). We detected 18 UFTLP 1.0 SNP loci in the SolCAP array (Supplemental Table S1, Supplemental Fig. S1), with 465 UFTLP 1.0 alleles that agreed with SolCAP SNP calls, and 23 alleles that did not. Additionally, 16 alleles remained unresolved due to missing SolCAP SNP calls.

Discussion

To identify SNPs in an F₂ segregation population of large-fruited fresh-market tomato, we used a WGS technique with variant calling steps that provided reliable sequence information of the entire genome. Furthermore, to increase the accuracy

Table 1. Genetic map of the large-fruited fresh-market tomato.

Chromosome no.	Markers (no.)	Genetic size (cM)	Physical size (Mbp) ^z	Recombination rate (cM/Mbp)	Mean distance between flanking markers (cM)	Maximum distance between flanking markers (cM)	Genetic intervals (no.) ^y	Mean distance between flanking markers (kbp)	Potential non-collinear markers (no.) ^x
1	312	279.6	90.9	3.08	0.90	41.6	4	285.2	5 (1)
2	909	230.6	53.5	4.31	0.25	25.1	4	58.8	4 (3)
3	363	140.9	65.3	2.16	0.39	16.4	1	74.6	5 (1)
4	318	173.4	64.5	2.69	0.55	13.8	1	194.4	3 (2)
5	73	156.9	65.3	2.40	2.18	18.2	4	893.7	5 (1)
6	84	113.7	47.3	2.40	1.37	33.2	1	548.2	4 (2)
7	204	136.3	67.9	2.01	0.67	22.1	2	323.8	9 (3)
8	401	232.0	64.0	3.63	0.58	20.6	2	159.0	9 (3)
9	615	233.0	68.5	3.40	0.38	30.8	1	111.1	5 (1)
10	178	155.1	64.8	2.39	0.88	10.3	1	362.6	9 (3)
11	77	56.9	54.4	1.05	0.75	13.4	1	596.2	5 (1)
12	80	115.3	66.7	1.73	1.46	15.6	3	823.4	5 (1)
Average				2.62	0.86	21.76		369.3	
Total	3,614	2,023.6	773				25		68 (22)

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

^yDefined by distances larger than 10 cM between flanking markers.

^xCollinearity between the genetic and physical orders of markers. Numbers in parentheses represent segments (a segment represents a group of consecutive markers).

of SNP calls, we applied an approach that integrated the results from the high-coverage WGS data of two parental lines and the low-coverage data of the biparental segregating population.

The total map length of the large-fruited fresh-market tomato was estimated to be 2023.6 cM. Previous studies reported that the total length of tomato genetic maps ranged from 867 to 1669.9 cM (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 et al., 1992), but these maps were developed using a cross between the modern tomato species, *S. lycopersicum*, and the wild tomato species, including *Solanum pimpinellifolium* and *Solanum pennellii*. To our knowledge, a genetic map constructed from a cross between *S. lycopersicum* and *S. lycopersicum* has not been previously reported, thus precluding the comparison of recombination rates with independent data. Nevertheless, the reduced map length calculated using marker down-sampling implied that the limited number of samples in our study could have affected the map length (Liu, 1998; Ott, 1999; Terwilliger and Ott, 1994). In addition, misplaced or incorrectly ordered markers due to the complex nature of the genomes between tomato lines could have affected the map length (Liu, 1998; Ott, 1999; Terwilliger and Ott, 1994). The current tomato genome assembly was built on the Heinz 1706 processing tomato cultivar (The Tomato Genome Consortium, 2012) that belongs to another type of tomato fruit class other than the tomato lines used in this study. Fully sequenced genomes from mapping populations will improve the power to resolve marker orders. We cannot rule out the possibility that the concentration of SNPs in short regions of the chromosomes and/or incomplete SNP calls in the population data (0.7% missing alleles in the final 3614 SNP sites) inflated the linkage map. However, we did not observe excessive distorted segregation patterns on specific chromosomes (data not shown), and a total of 1278 SNPs with distorted segregation were removed during genotypic error correction. Increasing the number of markers within distance intervals more than 10 cM between flanking markers could

increase the accuracy of the map distance estimation (Liu, 1998; Ott, 1999; Terwilliger and Ott, 1994).

Although strong recombination suppression was reported in the large pericentromeric regions within each tomato chromosome (Sim et al., 2012a; The Tomato Genome Consortium, 2012), we showed that higher rates of recombination occurred in the approximate centromere locations compared with these previous reports. We are aware that chromosome 11 of one of the parental lines in our study has a large-scale introgression between *S. lycopersicum* and *S. pimpinellifolium* (Alonge et al., 2020; S.F. Hutton, personal communication). Unlike other chromosomes, chromosome 11 showed the lowest recombination rate, and the genetic size was shorter than that of a previous SNP-based tomato map (Sim et al., 2012a). This suggests that the cross between the two more closely related parental lines in this study (i.e., two *S. lycopersicum* breeding lines sharing the same fruit class) created a higher level of recombination than that of a cross between *S. lycopersicum* and wild species. This result in turn provides evidence that the increased recombination rate in previously unidentified regions plays a role in determining the increased map length, at least in our study.

The mean 38% polymorphic rate for a complete comparison among core breeding lines in the UFTLP 1.0 provided intriguing evidence of DNA sequence diversity in tomato, especially in U.S. large-fruited fresh-market tomato. As a resource to the scientific community, we have provided the full allelic information of 58 core lines in the UFTLP 1.0, and information about allelic agreement between the UFTLP and SolCAP array SNP sites (Supplemental Table S1). Recently, several draft tomato sequence assemblies from long-read data were released (Alonge et al., 2020; Mazo-Molina et al., 2019; Schmidt et al., 2017; Wang et al., 2020). High-density linkage analysis, together with long sequence reads data and DNA insert libraries, were used to order and orient draft sequence assemblies into pseudo-molecule chromosomes (Ghurye and Pop, 2019; Rice and Green, 2019). Therefore, the genotypic dataset prepared for

Table 2. Distribution of the minor allele frequencies in fresh-market tomato linkage panel 1.0.

Minor allele frequency class	Markers (no.)	Proportion
0–0.1	788	0.22
0.1–0.2	292	0.08
0.2–0.3	478	0.13
0.3–0.4	1,438	0.40
0.4–0.5	618	0.17
Total	3,614	1.00

our study can be used to construct tomato consensus map(s) and understand tomato genome complexity. Having identified “common SNP sites” (i.e., UFTLP 1.0 SNP sites in agreement with the SolCAP array), the tomato community can revisit the array to improve its marker coverage and enhance the data-point acquisition time and affordability.

The large-fruited fresh-market tomato material used for this study represents a unique type of tomato fruit class, of which many lines have been successfully commercialized [e.g., ‘Tasti-Lee F₁’ (Bejo, Seeds, Oceano, CA), Scott et al., 2008] and transferred to public/commercial tomato research programs. A report released by the Florida Foundation Seed

Producers (2020) in 2018 revealed that 27 tomato lines currently generate revenue royalties. Although tomato is increasingly the target for genetic improvement, tomato researchers want to domestically and internationally expand the diversity of genetics and breeding resources. Because the UFTLP 1.0 has 20 of the aforementioned 27 lines and 38 new lines based on their current widespread use, and because a large proportion of SNP markers in the UFTLP 1.0 is expected to be polymorphic between any two lines, the UFTLP 1.0 can be used for tomato research.

Conclusions

This is the first study that we are aware of that uses an F₂ segregating population derived from a cross between two large-fruited fresh-market tomatoes to create a linkage map for this important horticultural crop. A genetic map of 3614 markers was constructed using WGS-based SNP genotyping, thereby generating the first version of a fresh-market tomato linkage panel (UFTLP 1.0). The genetic map and linkage panel will be important resources for future tomato research, including identifying associations between segregating alleles and phenotypes, and tomato biology.

Fig. 2. Recombination rates of the large-fruited fresh-market tomato genome. The approximate centromere locations are depicted by gray color. Ch. = chromosome.

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, doi: 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, doi: 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, doi: 0.1139/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, doi: 10.1093/oxfordjournals.jhered.a106504.
- Butler, L. 1936. Inherited characters in the tomato. II. Jointless pedicel. *J. Hered.* 27:25–26, doi: 10.1093/oxfordjournals.jhered.a104134.
- Cheema, J.J.S., T.H.N. Ellis, and J. Dicks. 2008. Exploring genome wide genetic linkage manifolds. In: *Proc. Genome Informatics Conf.* p. 36. (abstr.). 3 Dec. 2020. <<http://cbr.jic.ac.uk/dicks/software/threadmapper>>.
- 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 Vicente, M.C. and S.D. Tanksley. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134:585–596.
- Food and Agriculture Organization of the United Nations. 2016. Value of agricultural production. 3 Dec. 2020. <<http://www.fao.org/faostat/en/#data/QC/metadata>>.
- 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, doi: 10.1093/nar/gku1195.
- Florida Foundation Seed Producers. 2020. Florida Foundation Seed Producers, Inc. 3 Dec. 2020. <www.ffsp.net>.
- Florida Tomato Exchange. 2020. Florida Tomato Exchange. 3 Dec. 2020. <<https://www.floridatomatoes.org>>.
- Ghurye, J. and M. Pop. 2019. Modern technologies and algorithms for scaffolding assembled genomes. *PLOS Comput. Biol.* 15:e1006994, doi: 10.1371/journal.pcbi.1006994.
- 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, doi: 10.3835/plantgenome2018.02.0010.
- Guan, Z., T. Biswas, and F. Wu. 2017. The US tomato industry: An overview of production and trade. *Univ. Florida, Inst. Food Agr. Sci. FE1027*.
- Hudson, M.E. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Resour.* 8:3–17, doi: 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* 9:196, doi: 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, doi: 10.21273/JASHS04423-18.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079, doi: 10.1093/bioinformatics/btp352.
- Li, H. and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595, doi: 10.1093/bioinformatics/btp698.
- Liu, B. 1998. Statistical genomics: Linkage, mapping and QTL analysis. CRC Press, Boca Raton, FL.
- 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, doi: 10.1094/MPMI-01-19-0018-R.
- Ott, J. 1999. Analysis of human genetic linkage. Johns Hopkins Univ. Press, Baltimore, MD.
- Rice, E.S. and R.E. Green. 2019. New approaches for genome assembly and scaffolding. *Annu. Rev. Anim. Biosci.* 7:17–40, doi: 10.1146/annurev-animal-020518-115344.
- 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, doi: 10.1093/jxb/erq367.
- 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, doi: 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, doi: 10.21273/HORTSCI.43.7.2228.
- 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, doi: 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, doi: 10.1371/journal.pone.0045520.
- Stam, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *Plant J.* 3:739–744, doi: 10.1111/j.1365-3113.1993.00739.x.
- 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.
- Taylor, J. and D. Butler. 2018. Package ‘ASMap’ 3 Dec. 2020. <<https://cran.r-project.org/web/packages/ASMap/ASMap.pdf>>.

- Terwilliger, J.D. and J. Ott. 1994. Handbook of human genetic linkage. Johns Hopkins Univ. Press, Baltimore, MD.
- The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635–641, doi:10.1038/nature11119.
- University of Florida. 2020. HiPerGator. 3 Dec. 2020. <<https://www.rc.ufl.edu/services/hipergator>>.
- University of Florida, Institute of Food and Agricultural Sciences. 2020. UF/IFAS tomato breeding program. 3 Dec. 2020. <<https://tombreeding.ifas.ufl.edu>>.
- U.S. Department of Agriculture. 2005. Grades of tomatoes. 3 Dec. 2020. <<https://www.ams.usda.gov/grades-standards/tomato-grades-and-standards>>.
- U.S. Department of Agriculture. 2016. Tomatoes. 3 Dec. 2020. <www.ers.usda.gov/topics/crops/vegetables-pulses/tomatoes>.
- Wang, X., L. Gao, C. Jiao, S. Stravrovdis, 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, doi: 10.1038/s41467-020-19682-0.