Evaluation of Breeding Lines Derived from Wild *Solanum habrochaites* and a High-Throughput Phenotyping Multispectral-Imaging Robot for Improvement of Cultivated Tomato for Water-Limited Production Environments

By

BRYCE A. KUBOND

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Horticulture and Agronomy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Dina St. Clair, Chair

E. Charles Brummer

Allen Van Deynze

Committee in Charge

2023

Table of Contents:

Key Words:

Tomato, Plant Breeding, Climate Change, Water-Stress, Fruit Quality, *Solanum habrochaites*, Tomato Wild Relative, Bin Mapping, Combining Ability, High-Throughput Phenotyping

Abstract:

Climate change is expected to decrease precipitation in arid California, limiting tomato production in the state. Cultivated tomato, *Solanum lycopersicum*, has a limited genetic base for improvement. *S. habrochaites*, a tomato wild relative, is water-stress tolerant and may serve as a genetic source of abiotic stress-tolerance traits. With the advent of inexpensive and rapid genotyping platforms, high-throughput genomic methods have increasingly been developed to improve the speed and efficiency of plant breeding programs. These techniques are limited due to the lack of high-quality phenotype data to accompany the abundant genomic data available. High-throughput phenotyping (HTP) technologies are being developed to effectively harness high-throughput genomic methods to accelerate plant breeding for tolerance to water-limited environments.

Three experiments were performed in the field to evaluate introgression lines derived from *S. habrochaites* for their potential for improving cultivated tomato, and to determine the effectiveness of a HTP multispectral imaging robot for use in tomato breeding. Each of the three experiments utilized a split-plot experimental design, with main-plots each assigned to one of two water treatments, full crop evapotranspiration (ETc) or 40% ETc applied post-fruit set, and subplots assigned to different tomato genotypes. The first chapter evaluated a set of 24 introgression lines (ILs) derived from *S. habrochaites* and their recurrent parent and the data

were used to perform bin mapping for a set of 15 traits. The second chapter used 5 ILs from the first chapter, plus 3 inbred lines to obtain a set of 15 F1 hybrids using a North Carolina Design II mating scheme. These parents and their hybrids were evaluated for combining ability for a set of 11 traits. The third chapter used two sets of tomato genotypes, a training set and a validation set, which included ILs and modern hybrid cultivars to evaluate the effectiveness of an HTP robot at collecting accurate phenotype data for a set of 11 traits.

Breeding lines derived from *S. habrochaites* and HTP technologies have the potential to improve cultivated tomato. A total of 268 trait-genomic region associations (TGRAs) were identified among 22 of 24 ILs included in the first chapter. TGRAs were identified for each of the 15 traits. Horticulturally desirable TGRAs were identified for soluble solids content, fruit weight, degree of fruit sunburn, canopy cover, and maturity. IL LA3933 possessing an introgression on chromosome 4 from *S. habrochaites* may be suitable for use in developing inbred lines for a hybrid breeding program. GCA estimates of -52.55 to 75.21 were obtained for each of the 5 IL parents and 3 inbred line parents for each of the traits. Red pixel number data were successfully extracted from images collected by the HTP robot on a subset of experimental plots. A Spearman rank correlation of r = 0.76 was identified between manually collected ripe yield data and red pixel number, indicating the presence of a strong correlation. Additional data processing and analysis, possibly including the use of big data methodologies would be required to fully determine the potential of the HTP multispectral-imaging robot for tomato breeding.

Introduction:

Global climate change is impacting crop yields as temperatures increase. Increasing temperatures are expected to decrease precipitation in the arid western United States (Mote et al., 2005; Mahoney et al., 2021). Mountain snowpack in the western United States has been decreasing from the 1940s to present and is expected to continue to decline with increasing global temperatures (Mote et al., 2005). In California, climate change is predicted to decrease crop yields due to reduced precipitation and increasing temperatures (Lee et al., 2010; Diffenbaugh et al., 2015; Pathak et al., 2018; Ray et al., 2020).

Production of tomatoes (*Solanum lycopersicum*) in California relies on irrigation and is threatened by decreasing precipitation and water supply (Hartz et al., 2008; Pathak et al., 2018; Ray et al., 2020). California has a Mediterranean climate characterized by cool wet winters and dry warm summers (Deitch et al., 2017; Pathak et al., 2018; Seager et al., 2019). The California Central Valley receives little to no precipitation during the summer growing season of tomato. As a result, winter snowpack in the Sierra Nevada mountains is a key source of surface water for irrigation in the California Central Valley.

Tomato is the second most economically important vegetable crop worldwide behind potato (*Solanum tuberosum*) (FAOSTAT, 2021). In the United States, tomato production ranks second after potato among vegetable crops, with 274,000 acres and 912,000 acres planted in 2021, respectively (USDA NASS, 2022). California is the top producer of tomatoes in the United States with a total farmgate value of harvested tomatoes of 1.18 billion dollars in 2021 (USDA NASS, 2022). Two primary market classes of tomato are grown in California, fresh-market and processing (USDA NASS, 2022). Of the two market classes, processing tomatoes are the more

economically important, with a farmgate value of 905 million dollars in 2021 (USDA NASS, 2022). Over 95% of all processing tomatoes in the United States are produced in California (USDA NASS, 2022).

Cultivated tomato has a limited genetic base due to genetic bottleneck events that occurred during domestication (Rick, 1983; Miller and Tanksley, 1990; Corrado et al., 2013; Kulus, 2018; Tamburino et al., 2020). Tomato wild relatives, in contrast to cultivated tomato, are an excellent source of genetic diversity for expanding the limited genetic base of tomato for breeding (Rick, 1983; Miller and Tanksley, 1990). Cultivated tomato and its wild relatives possess a high degree of synteny and are inter-crossable with each other (Chetelat and Ji, 2007; Moyle, 2008). Previous tomato breeding efforts have used tomato wild relatives as sources of important traits, including both biotic and abiotic stress tolerances (Rick, 1983; Bai et al., 2018; Schouten et al., 2019). Multiple tomato wild relatives, including *Solanum habrochaites*, have been found to be tolerant to water-stress (Rick 1973; Rick, 1983; Spooner et al., 2005; Chetelat et al., 2009; Moyle and Muir, 2010; Dariva et al., 2020). The St. Clair lab has previously mapped QTL for water-stress tolerance traits to chromosome 9 of *S. habrochaites* (Truco et al., 2000; Goodstal et al., 2005; Arms et al., 2015; Lounsbery et al., 2016; Groh et al., 2022).

In addition to water-stress tolerance-related traits, *S. habrochaites* has been identified as a source of other beneficial traits for breeding cultivated tomato, including tolerance to root chilling, tomato leaf curl virus, and insect herbivores (Rick, 1973; Truco et al., 2000; Glas et al., 2012; Yang et al., 2014). Certain introgression lines (ILs) derived from *S. habrochaites* have been found to have increased soluble solids content in fruit compared to the cultivated recurrent parent (Bernacchi et al., 1998).

Plant breeders use IL libraries as useful resource for trait discovery and mapping of agriculturally important traits (Zamir, 2001). An IL library is a set of genotypes that each contain a different unique introgression from a wild species donor parent in the genetic background of a cultivated recurrent inbred parent. IL libraries have been created for multiple tomato wild relatives including *S. pennellii, S. neorickii, S. habrochaites*, *S. lycopersicoides*, and *S. sitiens* (Eshed and Zamir, 1994(a); Fulton et al., 2000; Monforte and Tanksley, 2000; Canady et al., 2005; Chetelat et al., 2019). Bin mapping uses IL libraries to map traits to a specific chromosomal region or bin (Eshed and Zamir, 1994(b); Chetelat et al., 2019). This form of trait mapping can be used to evaluate large portions of a donor parent's genomes with comparatively few resources due to the large spans of chromosome possessed by each IL.

F1 hybrid cultivars are the most common type of cultivar used in tomato production in the United States, similar to many other vegetable crops (Janick, 1998; Wehner 1999; Hartz et al., 2008; Processing Tomato Advisory Board, 2021). To develop F1 hybrid cultivars, inbred parental lines are created and then pairs of these lines are crossed with each other through controlled pollinations to obtain F1 hybrid(Fehr, 1987; Rehman et al., 2021). The F1 hybrid are then evaluated for traits of importance and compared to each other to determine the best combination of inbred parents that produce agriculturally superior performing F1 hybrids. Evaluation of combining ability is a way to assess and select inbred parents for producing the best performing F1 hybrid progeny.

High throughput genotyping methods with genomics tools have been developed to help increase the speed and efficiency of the plant breeding process (Bernardo 2008; Walsh 2009). Major limiting factors in utilizing high throughput genotyping methods for plant breeding are

the time and expense of obtaining high quality phenotype data for important traits on many individual plants or plots (Bernardo 2008; Walsh 2009; Cobb et al., 2013; Araus and Cairns, 2014). In order to address the gap between the amount of available genotype and phenotype data, high throughput phenotyping (HTP) methods are increasingly being developed and tested for use in plant breeding (Araus and Cairns 2014; Fahlgren et al., 2015). Developing HTP methods in order to fully capitalize on high throughput genomic methods for genotyping could be useful in accelerating the plant breeding process to address crop improvement for a rapidly changing climate (Furbank and Tester, 2011).

Objectives:

The overall objectives of my dissertation were to evaluate breeding lines derived from *S. habrochaites* introgressions in S. lycopersicum as a source of water-stress tolerance-related traits in breeding and to test a HTP multi spectral imaging robot for use in the improvement of water-stress tolerance in cultivated tomato.

The first chapter uses a subset of 24 introgression lines (Ils) from a *S. habrochaites* IL library (Monforte and Tanksley, 2000) to conduct bin mapping for a set of 15 traits including water-stress tolerance-related, horticultural, and fruit quality traits. Two water treatments were applied using a split-plot experimental design with main plots receiving one of two water treatments (normal and reduced). The ILs were assigned to the subplots. The objectives were to determine if there were significant differences for each trait of interest among this set of Ils and their recurrent parent, and to use bin mapping to test for the presence of statistically significant associations of water-stress tolerance-related, horticultural, and fruit quality traits to *S. habrochaites* genomic regions contained in the Ils.

The second chapter used the top performing Ils from the first chapter and a set of older inbred linesto create F1 hybrids using a Design II mating scheme. These hybrids, their parents, and the IL recurrent parent were evaluated for combining ability for 11 traits across 3 field locations. Two water treatments were applied to the genotypes using a split plot experimental design with main plots receiving one of two water treatments (normal and reduced). The objective of this study was to determine which IL parents from chapter 1 could be suitable for potential use in a F1 hybrid cultivar breeding program for the improvement of horticultural, fruit quality, and water-stress tolerance-related traits in processing tomato.

Chapter 3 concerns the evaluation of a tractor-based HTP system developed by Dr. David Slaughter at UC Davis for potential use in tomato breeding. Two sets of tomato genotypes, a training set and a validation set, were evaluated for 10 traits in replicated field trials using a split plot experimental design with two water treatments. Phenotype data was collected both manually and by the HTP robot phenotyping system for comparison. The objective of this chapter was to determine what (if any) relationship exists between manually collected phenotype data and data obtained by processing images or other sensor output collected by a tractor-based HTP system for ripe fruit yield.

Introduction References:

Araus, J.L. and Cairns, J.E. 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Frontiers in Plant Science.* 19(1): 52-61. https://doi.org/10.1016/j.tplants.2013.09.008

Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2015. High-resolution mapping of a major effect QTL from wild tomato *Solanum habrochaites* that influences water relations under root chilling. *Theoretical and Applied Genetics*. 128: 1713-1724. https://doi.org/10.1007/s00122-015-2540-y

Bernacchi, D., Beck-Bunn, T., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley S.D. 1998. Advanced backross QTL analysis in tomato. I. identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theoretical and Applied Genetics*. 97: 381-397. https://doi.org/10.1007/s001220050908

Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science*. 48(5): 1649-1664. https://doi.org/10.2135/cropsci2008.03.0131

Canady, M.A., Meglic, V. and Chetelat, R.T. 2005. A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome*. 48: 685-697. https://doi.org/10.1139/g05-032

Chetelat, R.T., and Ji, Y. 2007. Cytogenetics and evolution. In *Genetic Improvement of Solanaceous Crops* (Razdan, M.K. and Mattoo, A.K., eds). Science Publishers. Enfield, NH. Pp 77-112.

Chetelat, R.T., Pertuze, R.A., Faundez, L., Graham, E.B. and Jones, C.M. 2009. Distribution, ecology, and reproductive biology of wild tomatoes and related nightshades from the Atacama Desert region of northern Chile. *Euphytica*, 167: 77-93. https://doi.org/10.1007/s10681-008-9863-6

Chetelat, R.T., Qin, X., Tan, M., Burkart-Waco, D., Moritama, Y., Huo, X., Wills, T. and Pertuzé R. 2019. Introgression lines of *Solanum sitiens*, a wild nightshade of the Atacama Desert, in the genome of cultivated tomato. *The Plant Journal*. 100(4): 836-850. https://doi.org/10.1111/tpj.14460

Cobb, J.N., DeClerck, G. Greenburg, A., Clark, R. and McCouch, S. 2013. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phentoype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*. 126: 867-887. https://doi.org/10.1007/s00122-013-2066-0

Corrado, G., Piffanelii, P., Caramente, M., Coppola, M. and Rao, R. 2013. SNP genotyping reveals genetic diversity between cultivated landraces and contemporary varieties of tomato. *BMC Genomics*. 14: 835. https://doi.org/10.1186/1471-2164-14-835

Dariva, F.D., Copati, M.G.F., Pessoa, H.P., Alves, F.M., de Oliveira Dias, F., de Toledo Picoli, E.A., da Cunha, F.F. and Nick, C. 2020. Evaluation of anatomical and physiological traits of *Solanum*

pennellii Cor. Associated with plant yield in tomato plants under water-limited conditions. *Scientific reports*, *10*(1): 1-13. https://doi.org/10.1038/s41598-020-73004-4

Deitch, M.J., Sapundjieff, M.J. and Feirer, S.T. 2017. Characterizing precipitation variability and trends in world's Medditerranean-climate areas. *Water*. 9(4): 1-21.

Diffenbaugh, N.S., Swain, D.L. and Touma D. 2015. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*. 112(13): 3931-3936. https://doi.org/10.1073/pnas.1422385112

Eshed, Y. and Zamir, D. 1994 (a). A genomic library of *Lycopersicon pennellii* in *L. esculentum*: a tool for fine mapping of genes. *Euphytica*, 79: 175-180. https://doi.org/10.1007/BF00022516

Eshed, Y., and Zamir, D. 1994 (b). Introgressions from *Lycopersicon pennellii* can improve the soluble-solids yield of tomato hybrids. *Theoretical and Applied Genetics*. *88*(6-7): 891-897. https://doi.org/10.1007/BF01254002

Fahlgren, N. Gehan, M.A. and Baxter, I. 2015. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*. 24: 93-99. https://doi.org/10.1016/j.pbi.2015.02.006

Fehr, W.R. 1987. *Principles of Cultivar Development, vol. 1: Theory and Technique*. Iowa State University Press. Ames, Iowa. Pages 428-434

Food and Agriculture Organization of the United Nations. 2021. *FAOSTAT Crops*. FAOSTAT, Rome, Italy. http://www.fao.org/faostat/en/#data/QC (accessed 5 May. 2022).

Fulton, T.M., Grandillo, S., Beck-Bunn, T., Fridman, E., Frampton, A., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley, S.D. 2000. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics*. 100: 1025-1042. https://doi.org/10.1007/s001220051384

Furbank, R.T. and Tester, M. 2011. Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*. 16(12): 1360-1385. https://doi.org/10.1016/j.tplants.2011.09.005

Glas, J.J., Schimmel, B.C.J., Alba, J.M., Escobar-Bravo, R., Schurrink, R.C. and Kant, M.R. 2012. Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*. 13(12): 17077-17103. https://doi.org/10.3390/ijms131217077

Goodstal F.J., Kohler G.R., Randall L.B., Bloom A.J. and St. Clair D.A. 2005. A major QTL introgressed from wild *Lycopersicon hirsutum* confers chilling tolerance to cultivated tomato (*Lycopersicon escultentum*). *Theoretical and Applied Genetics*, 111: 898-905. https://doi.org/10.1007/s00122-005-0015-2

Groh, A.M., Kubond, B.A., and St. Clair, D.A. 2022. Fine mapping of QTL for water use efficiency-related traits on chromosome 9 of *Solanum habrochaites* in the field. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20828

Hartz, T.K., Miyao, G., Mickler., Lestrange, M., Stoddard., Nuñez, J., and Aegerter B. 2008. Processing Tomato Production in California. *University of California Agriculture and Natural Resources*, doi: 10.3733/ucanr.7228

Janick, J. 1998. Hybrids in horticultural crops. In K.R. Lamkey, and J.E. Staub (Ed.). *Concepts and Breeding of Heterosis in Crop Plants*. Pages: 45-56. Madison, Wisconsin, USA: CSSA Special Publication 25. https://doi.org/10.2135/cssaspecpub25.c4

Kulus, D. 2018. Genetic resources and selected conservation methods of tomato. *Journal of Applied Botany and Food Quality.* 91: 135-144. https://doi.org/10.5073/JABFQ/2018.091.019

Lee, J., De Gryze, S., Six, J. 2010. Effect of climate change on field crop production in California's Central Valley. *Climate Change*. 109: 335-353. https://doi.org/10.1007/s10584-011-0305-4

Lounsbery, J.K., Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2016. Quantitative Trait Loci for water-stress tolerance traits localize on chromosome 9 of wild tomato. *Crop Science*, 56: 1514-1525. https://doi.org/10.2135/cropsci2015.07.0432

Mahoney, K., Scott, J.D., Alexander, M., McCray, R., Hughes, M., Swales, D. and Bukovsky, M. 2021. Cool season precipitation projections for California and the Western United States in NA-CORDEX models. *Climate Dynamics*. 56: 3081-3102. https://doi.org/10.1007/s00382-021-05632-z

Miller, J.C. and Tanksley, S.D. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics*. 80: 437-448. https://doi.org/10.1007/BF00226743

Monforte, A.J., and Tanksley, S.D. 2000. Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicum hirsutum* genome in a *L. esculentum* genetic background: a tool for gene mapping and gene discovery. *Genome*, 43(5): 803-813. https://doi.org/10.1139/g00-043

Mote, P.W., Hamlet, A.F., Clark, M.P. and Lettenmaier, D.P. 2005. Declining mountain snowpack in western North America. *Bulletin of the American Meterological Society*. 86(1):39-49. https://doi.org/10.1175/BAMS-86-1-39

Moyle, L.C. 2008. Ecological and evolutionary genomics in the wild tomatoes (Solanum sect. Lycopersicon). *Evolution*. 62(12): 2995-3013. https://doi.org/10.1111/j.1558-5646.2008.00487.x

Moyle L.C. and Muir C.D. 2010. Reciprocal insights into adaptation from agricultural and evolutionary studies in tomato. *Evolutionary Applications*. 3(5-6): 409-421. https://doi.org/10.1111/j.1752-4571.2010.00143.x

National Agricultural Statistics Service. 2021. *Statistics by subject*. USDA, Washington DC. http://www.nass.usda.gov/Statistics_by_Subject/?sector=CROPS (accessed 5 May 2022).

Pathak, T., Maksey, M.L., Dahlberg, J.A., Kearns, F., Balie, K.M. and Zaccaria, D. 2018. Climate change trends and impacts on California agriculture: a detailed review. *Agronomy*. 8(25): 1-27. https://doi.org/10.3390/agronomy8030025

Processing Tomato Advisory Board. 2021. *2021 Top 50 Variety Report*. PTAB, Davis, CA. http://www.ptab.org/2021Top50.pdf (accessed 7 September 2022)

Ray, P., Wi, S., Schwarz, A., Correa, M., He, M.X. and Brown, C. 2020. Vulnerability and risk: climate change and water supply from California's central valley water system. *Climate Change*. 161(1): 177-199. https://doi.org/10.1007/s10584-020-02655-z

Rehman, U.A., Dang, T., Qamar, S., Ilyas, A., Fatema, R., Kafle, M., Hussain, Z., Masood, S., Iqbal, S. and Shahzad, K. 2021. Revisiting plant heterosis – from field scale to molecules. *Genes*. 12(11): 1688. https://doi.org/10.3390/genes12111688

Rick C.M. 1973. Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. *Genes, Enzymes, and Populations*. 255-269. https://doi.org/10.1007/978-1-4684-2880-3_17

Rick C.M. 1983. Genetic variability in tomato species. *Plant Mol. Biol. Rep.* 1: 81-87. https://doi.org/10.1007/BF02680303

Seager, R., Osborn, T.J., Kushnir, Y., Simpson, I.R., Nakamura, J. and Liu, H.B. 2019. Climate variability and change of Mediterranean-type climates. *Journal of Climate*. 32(10): 2887-2915. https://doi.org/10.1175/JCLI-D-18-0472.1

Spooner, D.M., Peralta, E. and Knapp, S. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes (*Solanum L.* section *Lycopersicon* (Mill.) Wettst.). *Taxon*. 54:43-61.47. https://doi.org/10.2307/25065301

Tamburino, R., Sannino, L., Cafasso, D., Cantarella, C., Orru, L., Cardi, T., Cozzolino, S., D'Agostino, N. and Scotti, N. 2020. Cultivated tomato (*Solanum lycopersicum* L.) suffered a severe cytoplasmic bottleneck during domestication: implications from chloroplast genomes. *Plants-Basel*. 9(11): https://doi.org/10.3390/plants9111443

Truco, M.J., Randall, L.B., Bloom A.J. and St. Clair D.A. 2000. Detection of QTLs associated with shoot wilting and root ammonium uptake under shilling temperatures in an interspecific backcross population from *Lycopersicum esculentum* x *L. hirsutum*. *Theoretical and Applied Genetics*. 101:1092-1092. https://doi.org/10.1007/s001220051583

Walsh, B. 2009. Quantitative genetics, version 3.0: where have we gone since 1987 and where are we headed? *Genetica*. 136(2): 213-223. https://doi.org/10.1007/s10709-008-9324-0

Wehner, T.C. 1999. Heterosis in vegetable crops. *Genetics and Exploitation of Heterosis in Crops*. In J.G. Coors and S. Pandey (Ed.). *Genetics and Exploitation of Heterosis in Crops*. Pages: 387-397. Madison, Wisconsin, USA: American Society of Agronomy, Crop Science Society of America

Yang, X.H., Caro, M., Hutton, S.F., Scott, J.W., Guo, Y.M., Wang, X.X., Rashid, M.H., Szinay, D., de Jong, H., Visser, R.G.F., Bai, Y.L. and Du, Y.C. 2014. Fine mapping of the tomato yellow leaf curl virus resistance gene Ty-2 on chromosome 11 of tomato. *Molecular Breeding.* 34(2): 749-760. https://doi.org/10.1007/s11032-014-0072-9

Zamir, D. 2001. Improving plant breeding with exotic genetic libraries. *Nature Reviews*. 2: 983-989. https://doi.org/10.1007/s11032-014-0072-9

Chapter 1 (As Appears for Early Access in *Crop Science*):

BIN MAPPING OF WATER STRESS TOLERANCE-RELATED, FRUIT QUALITY, AND HORTICULTURAL TRAITS IN TOMATO INTROGRESSION LINES DERIVED FROM WILD *SOLANUM HABROCHAITES*

1. Core Ideas:

- Use of *S. habrochaites* as a source of genetic diversity for breeding improvement of cultivated tomato
- Genotype by environment interactions were present for each trait studied
- TGRAs for increased water-stress tolerance traits were not detected
- TGRAs were discovered that could be used to improve certain horticultural and fruit quality traits

2. Authors:

Bryce A. Kubond* and Dina A. St. Clair

Affiliations: Plant Sciences Department, University of California Davis, Asmundson Hall, Mail Stop 3, One Shields Ave., Davis, CA, USA 95616. Received:_____. *Corresponding author (bakubond@ucdavis.edu).

3. Abstract

Climate change is reducing water availability for crop production. Cultivated tomato, *Solanum lycopersicum,* has a limited genetic base for improvement. A wild tomato, *Solanum habrochaites,* is water-stress tolerant and may serve as a genetic source of abiotic stress-tolerance traits. A set of 24 introgression lines (ILs) derived from *S. habrochaites* and processing

tomato inbred cultivar E6203 were evaluated in the field in replicated experiments over three years for 15 traits, including water-stress tolerance-related, fruit quality, and horticultural traits. A split-plot experimental design was employed with reduced and full irrigation treatments as main plots. Subplots consisted of the ILs and control E6203, and trait data was collected on a per subplot basis. Statistical data analyses and bin mapping were performed on a per trait basis. Trait-genomic region associations (TGRAs) were detected in ILs when a significant difference between E6203 and an IL was present. A total of 268 TGRAs were detected for all 15 traits. Traits mapped to introgressions in 22 of the 24 ILs and on 11 of 12 tomato chromosomes. ILs contributed both positive and negative allelic effects at TGRAs. TGRA with positive allelic effects from *S. habrochaites* were identified for soluble solids content, fruit weight, degree of fruit sunburn, canopy cover, and maturity. Our results suggest that *S. habrochaites* may be a useful resource for breeding improvement of certain fruit quality and horticultural traits in cultivated tomato. Overall, *S. habrochaites* alleles did not contribute to increased water-stress tolerance relative to E6203 at the milder level of water-stress used in this study.

4. Abbreviations:

BioM, above ground plant fresh biomass; CRT, total carotenoids content; DAPG, days after transplant to first green fruit of 1 cm diameter or larger; DAPR, days after transplant to first ripe fruit; ETc, crop evapotranspiration; FW, average fruit weight; HI, harvest index; ILs, introgression lines; QTL, quantitative trait locus or loci; RYLD, ripe yield; SDW, shoot dry weight; SFW, shoot fresh weight; TGRAs, trait-genomic region associations; TYLD, total yield

5. Introduction

Tomato (*Solanum lycopersicum*) is one of the most economically important vegetable crops grown globally, second only to potato (*Solanum tuberosum*) (FAOSTAT, 2021). In the United States, tomato production ranks second after potato (National Agricultural Statistics Service, 2021). California is the leading producer of tomatoes in the United States (National Agricultural Statistics Service, 2021). Tomatoes (both processing and fresh market) are one the most economically important crops in California, with a farm gate value of 1.19 billion US dollars in 2020, and over 95% of US processing tomatoes are produced in California (National Agricultural Statistics Service, 2021).

Global climate change is predicted to decrease yields of California tomatoes due to reduced precipitation and increasing temperatures over the next century (Lee et al., 2010; Diffenbaugh et al., 2015). California possesses a Mediterranean climate characterized by wet cool winters and dry warm summers (Diffenbaugh et al., 2015). As a result of California's dry summers, winter snowpack in the Sierra Nevada mountains is an important source of water for irrigation during the dry months. Climate change has led to a decrease the mountain snowpack in the arid western United States from the 1940s to the present and this decline is expected to continue with further increases in global temperatures (Mahoney et al., 2021; Mote et al., 2005). Decreasing precipitation and snowpack poses a threat to water resources needed to produce crops in California, including tomato which relies on irrigation, reducing the feasibility and sustainability of continued tomato production (Hartz et al., 2008).

Cultivated tomato has a limited genetic base due to multiple genetic bottleneck events during domestication (Rick, 1983). Fortunately, wild tomato species are a rich source of genetic diversity for expanding the genetic base of cultivated tomato for breeding improvement (Rick,

1983; Miller and Tanksley, 1990). Tomato wild relatives have been used in breeding as a source of valuable traits, including biotic and abiotic stress tolerances (Bai, et al., 2018; Rick, 1983; Schouten et al., 2019). Cultivated tomato and its wild relatives share a high degree of synteny and are inter-crossable with each other (Chetelat and Ji 2007; Moyle 2008). Wild *Solanum habrochaites* has been found to be water-stress tolerant (Rick, 1983; Spooner et al., 2005). In our lab, QTL for water-stress tolerance-related traits were mapped to chromosome 9 of *S. habrochaites* (Truco et al. 2000; Goodstal et al., 2005; Arms et al., 2015; Lounsbery et al., 2016). To the best of our knowledge, there are no other published studies reporting QTL in *S. habrochaites* for water-stress tolerance-related traits.

Introgression lines (ILs) derived from wild species are a useful tool for trait discovery and mapping agriculturally important traits (Zamir 2001). A set of ILs are inbred lines that contain the same recurrent parent genetic background and differ only by a unique chromosome introgression or introgressions transferred from a wild species donor parent (Zamir 2001). A set of *S. habrochaites* ILs (i.e., an IL library) in the recurrent parent background of inbred processing tomato cultivar E6203 was developed by Monforte and Tanksley (2000). IL libraries representing the majority of a wild species genome have been developed for other tomato wild relatives, including *S. pennellii, S. neorickii*, S*. lycopersicoides,* and *S. sitiens* (Eshed and Zamir, 1994(a); Fulton et al., 2000; Canady et al., 2005; Chetelat et al., 2019). IL libraries can be used to discover trait-genotypic region associations (TGRAs) using bin mapping (Eshed and Zamir, 1994(b); Chetelat et al., 2019). In bin mapping, trait phenotypic data for each IL is compared to that of its recurrent cultivated parent to determine if the presence of a specific wild introgression is significantly statistically associated with phenotypic variation in a given trait.

Because a specific IL and its recurrent parent genetically differ only in the chromosomal area of an introgression, significant statistical differences in the trait means between the two lines are attributed to the presence of the introgression. Overlapping introgressions in ILs create “bins” that can be used to further narrow the chromosomal location of region(s) containing one or more QTL controlling a trait.

Water-stress tolerance-related traits can be phenotypically evaluated on plants grown under deficit irrigation (Payero et al., 2009; Richards 2006; Tuberosa 2012). Deficit irrigation (in contrast to drought) is the continued application of water throughout a growing season below the total evapotranspiration needs of a particular crop. It is technically difficult and laborious to directly measure water-stress tolerance of plants in a field situation (Richards 2006). Alternatively, traits such as yield, harvest index, shoot biomass, and carbon isotope discrimination are commonly used to indirectly measure water-stress tolerance (Payero et al., 2009; Richards 2006; Tuberosa 2012). Traits used to indirectly measure water-stress tolerance that are positively correlated to water-stress tolerance include abscisic acid concentration, chlorophyll concentration, early vigor, harvest index, and yield (Payero et al., 2009; Richards 2006; Tuberosa 2012). Other traits used to indirectly measure water-stress tolerance that are negatively correlated with water-stress tolerance include carbon isotope discrimination and dry biomass (Richards 2006; Tuberosa 2012). These various indirect measurements have been used in a wide variety of crops including corn (Richards 2006; Payero et al., 2009; Tuberosa 2012), tomato (Lounsbery et al., 2016; Groh et al., 2022), wheat (Richards 2006; Tuberosa 2012), sorghum (Richards 2006), and soybean (Richards 2006; Tuberosa 2012). Previous studies of water-stress tolerance in tomato have employed yield, shoot biomass, specific leaf area, and

carbon isotope discrimination as indirect measurements (Lounsbery et al., 2016; Groh et al., 2022).

Key horticultural and fruit quality traits that are targeted for improvement by plant breeders in processing tomatoes include fruit maturity, yield, plant size, fruit soluble solids (Brix), pH, fruit size, fruit weight, and lycopene (a red carotenoid pigment) content (Barrett et al., 2007; Hartz et al., 2008). Tomato processors require tomatoes with a pH below 4.4 to maintain food safety (Anthon et al., 2011). Carotenoids, including lycopene, are antioxidants that provide human health benefits (Rao and Rao 2007). Processing tomato growers are compensated primarily based on fruit yield; therefore, it is a key trait of interest for improvement. Processing tomato plants must have a determinant, compact growth habit (i.e., branches terminate with inflorescences), and uniform maturity of fruit to permit efficient once-over machine harvest at the end of the growing season (Hartz et al., 2008).

In addition to water-stress tolerance-related traits, *S. habrochaites* possesses other agriculturally beneficial horticultural and fruit quality traits. *S. habrochaites* is tolerant to root chilling, tomato yellow leaf curl virus, and resistant to various herbivores including insects (Glas et al., 2012; Rick 1973; Truco et al., 2000; Yang et al., 2014). Additionally, some introgression lines from *S. habrochaites* exhibit increased soluble solids content in fruit when compared to the cultivated recurrent parent (Bernacchi et al., 1998(a)).

The objectives for this study were to: (1) determine if there were significant differences for each trait of interest among a set of ILs derived from *S. habrochaites* and their recurrent parent, and (2) test for the presence of statistically significant associations of fruit quality,

water-stress tolerance-related, and horticultural traits to *S. habrochaites* genomic regions in the ILs using bin mapping.

6. Materials and Methods

6.1 Plant Material

A set of tomato introgression lines (ILs) derived from wild species *S. habrochaites* accession LA1777 was used for this study (Table 1). Seeds of the IL accessions were obtained from the C.M. Rick Tomato Genetics Resource Center (TGRC) at UC Davis (http://tgrc.ucdavis.edu). Each IL in the set contains one or more unique introgressions for 11 of the 12 tomato chromosomes from *S. habrochaites* in the recurrent parent genetic background of *S. lycopersicum* processing tomato inbred cultivar E6203 (Monforte and Tanksley, 2000).

Several plants of each IL were grown during summer 2015 in 3L pots in a UC Davis greenhouse in Davis, California under natural light and drip irrigation to obtain sufficient amounts of self-pollinated seed to conduct replicated field experiments. Observations on plant fertility, fruit set, and plant habit (i.e., determinancy) for each IL were obtained. These traits are important to help ensure that each IL chosen for the study would most likely set fruit in the field and maintain a determinant growth habit, as is required for processing tomato cultivars. One IL, LA3945, containing a chromosome 6 introgression, was dropped from the study due to indeterminant (non-compact) growth habit. A total of 24 ILs were included in the subsequent field experiments (Table 1). These ILs represented approximately 60% of the genome of *S. habrochaites*. The control for this experiment was the recurrent parent of the ILs, *S. lycopersicum* inbred processing tomato cv. E6203.

6.2 Field Experimental Design

Field experiments were conducted during summers of 2016, 2017, and 2018 at the UC Davis Plant Sciences Field Research Facility in Davis, California. A location for this experiment was defined as fields with different cropping histories. A split plot experimental design was used, with each of two locations containing two repetitions of the split plot. Each split plot experiment consisted of two main plots. Main plot treatments were two different irrigation regimes (normal and reduced (deficit), see next paragraph). Main plots contained 4 blocks of subplots. Subplots were assigned genotypes (24 ILs and E6203 as control). Eight plants were included in each subplot. A single row of plants at 30.5 cm spacing between plants was planted on each 154.2 cm center bed. A within-row alley of 91.4 cm without plants was included between the end and beginning of each subplot in the row to facilitate access for data collection. Double border rows of various inbred processing tomato cultivars were placed between each main plot, as well as on the outside perimeter of the field experiment, to minimize edge effects.

Prior to field preparation, seeds of all ILs and control E6203 were seeded into flats in a UC Davis greenhouse during March of each year. Plants were grown until they reached the second true leaf stage (approximately 5 weeks). Flats were then transferred to a lathhouse to harden off for a week before transplanting to the field by hand.

After transplanting, seedlings were sprinkler irrigated for approximately 3 weeks to allow for root establishment prior to switching to subsurface drip irrigation. Water was delivered to the plants via 1.58 cm width drip tape with 30.48 cm emitter spacing (Toro Flow Control) subsurface drip irrigation, with a single line buried 20 cm to 22 cm below the soil

surface of each bed. Water was delivered 3 times weekly to evenly space the irrigation water application over weekly periods. The amount of water delivered during each irrigation was determined by the amount of evapotranspiration (ET), as recorded by a nearby CIMIS station (cimis.water.ca.gov) since the prior irrigation, and the canopy size of the growing tomato plants. The rate of water flow was calculated prior to each irrigation due to daily fluctuations in water pressure. The daily irrigation length for each treatment was calculated using the water flow rate of the irrigation system and the calculated amount of water to apply to each treatment as determined by the crop evapotranspiration (ETc) for tomato (see next paragraph). The irrigation system was manually turned on and off for each irrigation period. The amount of water applied to each treatment and location combination was recorded using water flow meters (Sensus VMSR2 Brass Water Meter).

Two irrigation treatments were used: normal or full crop evapotranspiration (ETc) of tomato for the duration of the season, and a reduction post-full fruit set to 40% of ETc for the remainder of the season. The reduced water treatment used in this experiment was selected due to the possibility of adoption by California processing tomato growers to save water if it does not lead to significant yield losses. Full fruit set was defined as 51% of E6203 control plots scored as having a first ripe fruit. After full fruit set was reached, the reduced water treatment was implemented post-fruit set to limit impacts on yield. Tomato floral abortion can be due to water-stress (Ruan et al., 2012). Deficit irrigation after fruit set is achieved limits the amount of yield loss from floral abortion. The amount of water applied to the reduced water treatment was determined by multiplying the amount of water to apply to the full water treatment by 0.4. All subsurface drip irrigation water was precisely applied and the amount was calculated based

on canopy width measurements and ETc. During the field experiments, no measurable summer precipitation was recorded. Urea nitrogen fertilizer was banded prior to transplant at a depth of 15cm to 20cm beneath the surface at a rate of 22.41 Kg/ha (Hartz 2008). Additional urea nitrogen was applied via fertigation for six consecutive weeks, with a weekly application rate of 22.41 Kg/ha starting 21 days post-transplant.

In 2017 and 2018, the same experimental design and field practices were utilized, except the number of plants in a subplot was increased from 8 to 10. In addition, urea nitrogen fertilizer was applied via fertigation for four weeks at a rate of 33.6 Kg/ha starting 21 days post-transplant instead of six weeks per revised recommendations by field research support staff.

6.3 Phenotyping Traits

In all three years, trait data was collected on a per-subplot basis. The traits were grouped into three categories: horticultural, fruit quality, and water-stress tolerance-related (Table 2). Horticultural and water-stress tolerance-related traits were collected in each year. Due to labor limitations, fruit quality traits were only collected in 2016 and 2017.

Days after planting to first green fruit (DAPG) was defined as the number of days from transplanting until greater than 50% of the plants in a subplot each had a green fruit of at least 1 cm in diameter. Scoring was conducted three times a week, on Mondays, Wednesdays, and Fridays. Days to first red fruit (DAPR) was defined as the number of days from transplanting until greater than 50% of the plants in a subplot had at least one fully ripe fruit (Lounsbery et al., 2016). As with DAPG, DAPR was scored three times weekly on Mondays, Wednesdays, and Fridays.

Plant growth habit, degree of fruit sunburn, and canopy cover were scored when each subplot was ready to be harvested (see next paragraph). Subplots were subjectively scored for habit on a scale of 1 to 5, ranging from very prostrate to very upright. Sunburn and Canopy were also scored on a subjective 1 to 5 scale, from no fruit sunburn to severe sunburn, and sparse canopy cover to dense canopy cover, respectively.

Subplots were destructively harvested when all the plants in a subplot had approximately 90-95% ripe fruit. If a subplot was later maturing relative to E6203, it was harvested at the end of the growing season regardless of percentage of ripe fruit load. Two plants from each subplot were cut at the soil line. Fruit was removed manually from the two shoots per subplot, separated into fully ripe and not ripe, and each group was weighed to obtain total fruit yield and ripe fruit yield. Subsequently, the two shoots were placed into an onion mesh bag and weighed immediately to obtain fresh shoot weight (Lounsbery et al., 2016). Harvest index was calculated using the following formula: HI = RYLD/(TYLD + Shoot Fresh Weight). Total fresh biomass was determined by adding total fruit yield to shoot fresh weight. Onion mesh bags containing the harvested tomato shoots were placed in a forced air dryer and allowed to dry for at least 2 weeks, then weighed to obtain shoot dry weight per subplot (Lounsbery et al., 2016).

Once the plants in each subplot had sufficient numbers of ripe fruit, based on of visual inspection of subplots, 25 ripe fruit were selected randomly and harvested. The 25 fruit were weighed and the weight divided by 25 to obtain the average fruit weight (FW) per subplot. After weighing, each fruit was cut in half longitudinally and one half from each of the 25 fruits were blended together for 1 minute. The resulting tomato slurry was measured immediately for

degree Brix using a digital hand-held pocket refractometer (Atago LTD) and for pH using a portable pH meter (Oakton pH 150). Subsequently, five aliquots of the tomato slurry from each subplot were sampled into 2 ml microcentrifuge tubes and frozen at -80 ˚C for later lab analysis of carotenoids.

Total carotenoids for each subplot was obtained using one 2 ml aliquot of tomato slurry that was thawed to room temperature. 100 µl of tomato slurry was added to 1.9 ml of a 2:1:1 hexane, acetone, ethyl acetate solution at room temperature (Laur and Tian, 2011). The solution was then sonicated for 10 minutes to extract the pigment from the slurry. After sonication the solution was centrifuged at 13,000 rpm for 2 minutes, and repeated as needed to remove any tomato slurry remaining in suspension. The solution without a suspended slurry was then read at 450nm with a UV spectrophotometer to obtain total carotenoids content (Laur and Tian, 2011).

6.4 Statistical Analysis

Statistical analysis was performed on a per trait basis. Analysis of variance (ANOVA) was performed for each trait using the R stats package (R Core Team) and lme4 (Bates et al., 2015). Each trait dataset was checked for normality with Shapiro-Wilk W-statistic and with a Quantile-Quantile plot (R Core Team). Homogeneity of variance was evaluated using residuals plots (R Core Team). Each trait (except HI16) met both ANOVA assumptions of normality and homogeneity of variance. HI16 required a logit transformation in order to meet both ANOVA assumptions.

Initially, trait data was pooled across the three years for analysis, except for fruit quality traits which were pooled across the two years for which data was collected (2017 and 2018).

Statistically significant ($P \leq 0.05$) year x genotype main effect interactions were detected for each trait. Due to statistically significant interactions between year and genotype main effects, subsequent ANOVA for each trait was performed separately for each year using the following linear additive model with the lmer function in the lme4 package in R (Bates et al., 2015):

Trait = Loc + Water + Genotype + (Water x Loc) + (Genotype x Loc) + (Genotype x Water) + (Genotype x Water x Loc)

Loc refers to field location (loc1 or 2) within each year, Water refers to water treatment (Full ETc or 40% ETc post-fruit set), Genotype refers to the 24 ILs and E6203 control. When significant ($P \leq 0.05$) genotype x environment (G x E) interactions were detected for a particular trait dataset, subsequent analysis was performed by year, location, treatment, or a combination, depending on the source of variation exhibited in an interaction.

After ANOVA was conducted for each trait dataset, mean separations were performed with Dunnett's Multiple Comparison (Dunnet's), which compared each IL to the control E6203 with the cld(contrasts) function in the lsmeans package (Lenth et al., 2016). When significant ($P \leq 0.05$) G x E interactions (genotype x location or genotype x water treatment) were detected, trait data is referred to by their specific year, location, and treatment combinations. For example, total fruit yield (TYLD) from location 1 under the full water treatment in 2016 is denoted as TYLD161F, with the year being appended after the trait code, either 16 for 2016 or 17 for 2017, followed by the location (1 or 2) and finally the water treatment, F for full water treatment or R for reduced water treatment.

Mean separations for each trait dataset were also performed using Tukey's, which compared all ILs plus E6203 to each other. The cld(contrasts) function in the lsmeans package

was used to obtain means for Tukey's (Lenth el al., 2016). Spearman rank correlations among pairs of traits were computed with the cor.test function in the stats package (R Core Team).

6.5 Bin Mapping

Bin mapping for the set of ILs was performed per trait after any significant genotype x environment interactions present were separated to resolve interactions between main effects. Dunnett's was used for mean comparisons between each IL and E6203 due to a higher number of mean comparisons than the number of *t*-tests used by Eshed and Zamir (1994(b)). The original significance threshold of $P \leq 0.05$ was adjusted to $P \leq 0.002$ using a Bonferroni correction to reduce type 1 error. When a significant difference between the trait means of an IL and E6203 was detected with Dunnett's, that trait was mapped to the chromosome introgression contained within the IL. This method was adapted from Eshed and Zamir (1994(b)). Very few of the ILs included in our study had overlapping chromosomal introgressions, unlike in Eshed and Zamir (1994(b)). As a result, trait locations were mapped to a particular IL's introgression and not to a more specific chromosomal bin. The ILs in this study that contain overlapping introgressions are as follows: LA3921, LA3922, and LA3965 overlap on chromosome 2; LA3953 and LA3955 overlap on chromosome 8; LA3957 and LA3956 overlap on chromosome 9; LA3958 and LA3960 overlap on chromosome 9; LA3960, LA3963, and LA3965 overlap on chromosome 10; and LA3960 and LA3969 overlap on chromosome 12 (Figure 2).

A trait genetic region association (TGRA) was declared when there was a significant statistical difference detected between trait means of an IL and E6203 with Dunnett's. A region with a significant TGRA can contain one or more QTL that influence a trait due to the large genetic size of the introgressions. Interpretations of trait performance were made relative to

horticultural desirability in terms of tomato as a crop, and not absolute values compared to E6203.

7. Results

7.1 ANOVA

Significant genotype x year main effect interactions ($P \leq 0.05$) were identified for all traits. Each trait was subsequently analyzed by year. If additional G x E interactions (e.g., interactions between genotype and year, location, or water treatment) were identified, traits were evaluated by location within each year and, in some cases, by water treatment at each location in each year. For all traits (except TYLD181F and BioM18F), genotypes were significantly different ($P \leq 0.05$) (Table 3). Significant genotype x water treatment interactions were detected for the following traits in at least one location in at least one year: TYLD, RYLD, SFW, SDW, BioM, FW, Brix, pH, and CRT. Horticultural traits not associated with water-stress tolerance (i.e., Canopy, Sunburn, Habit, DAPG, and DAPR) did not exhibit any genotype x treatment interactions after trait data was separated by year.

Significant genotype x location interactions were detected for all traits (except Sunburn and Canopy) in at least one year. Significant genotype x treatment interactions were present for all fruit quality traits, and for 5 out of 11 traits in the water-stress tolerance/horticultural trait categories (see Table 2) (Supplemental Table 1).

7.2 Means Separation

Significance differences ($P \leq 0.05$) were present among genotype means with Tukey's for each of the traits, except TYLD181F, BioM181F, and SDW17R (Supplemental Table 2). None of

the ILs had increased horticultural performance compared to E6203 for all or the majority of traits. Two ILs, LA3956 and LA3957, were not significantly different from E6203 for any traits.

Significant differences were detected between some ILs and E6203 with Dunnett's test. Overall, the ILs performed similarly or were inferior to E6203 horticulturally. At least one IL performed better horticulturally compared to E6203 for the following horticultural or fruit quality traits: Canopy, Sunburn, DAPR, DAPG, Brix, pH, and FW. None of ILs exhibited significantly superior horticultural performance compared to E6203 for any of the traits associated with water-stress tolerance. Two ILs, LA3956 and LA3957, that both contain an introgression on chromosome 9, were not significantly different than E6203 for any trait.

For the majority of traits, there were no significant main effect interactions between genotype and water treatment (Table 3). This indicates that the genotype means were not statistically significantly different in rank nor magnitude between the two water treatments. For the traits that exhibited genotype x water treatment main effect interactions, trait means for each genotype can be compared between the two treatments at a particular location. E6203 means for the traits in which significant genotype x water treatment main effects were detected were not consistently higher for the full water treatment (Supplemental Table 2). E6203 is not known to be tolerant to water-stress and therefore would be expected to have worse horticultural performance under water-stress than at full ETc. Despite some ILs pH values being significantly different from each other and E6203, all IL means (range 4.11 to 4.57) were within a commercially acceptable range.

Mechanical harvest requires processing tomato fruit to be within a certain size and weight range. All of the ILs and E6203 had relatively small fruit. *S. lycopersicum* cv. Peto 95-45 is a small fruited inbred processing tomato cultivar which has a minimum commercially acceptable fruit size for mechanical harvest (Hartman and St. Clair, 1999). Since each of the ILs and E6203 had fruit sizes that were smaller than Peto 95-43, they would be too small be considered commercially acceptable.

7.3 Spearman Rank Correlations

Significant ($P \leq 0.05$) correlations were detected at $|r| \geq 0.396$ and considered of biological and breeding interest. Interpretations of correlations were based on horticultural desirability of each trait, not on the direction of the correlation. In general, either positive ($|r| \geq 0.396$) or non-significant ($|r| \leq 0.395$) correlations were observed among horticultural and water-stress tolerance-related traits (Figure 1).

The traits Canopy, Sunburn, and Habit were analyzed separately from the remaining traits because they are categorical (Supplemental Table 3). For these traits, negative correlations were found between Canopy and Habit (range: $|r| = 0.41$ to 0.91). Sunburn and Canopy as well as Sunburn and Habit were correlated (range: $|r| = 0.56$ to 0.91).

In this study, fruit quality traits tended to show low to no correlations to other fruit quality traits. Despite this trend, when Brix was compared to HI, SFW, SDW, DAPG, and DAPR, significant negative correlations were obtained (range: $|r| = 0.40$ to 0.81). FW had some positive correlations to TYLD, RYLD, and HI (range: $|r| = 0.40$ to 0.73).

7.4 Bin Mapping

Statistically significant ($P \leq 0.002$, using Bonferroni correction) trait-genomic region associations (TGRAs) were discovered for 22 of the 24 introgression lines (Figure 2). The two ILs without any TGRAs were LA3956 and LA3957, both containing an introgression on chromosome 9. Trait-genomic region associations were detected for each of the 11 of the 12 chromosomes included in the study. Among the 22 ILs, 268 TGRAs were observed (Supplemental Table 4). Each trait was mapped to at least one IL in at least one of the three years of experiments. In general, TGRAs in the ILs had an undesirable allelic effect horticulturally. However, some TGRAs with positive allelic effects from *S. habrochaites* were detected for traits in the horticultural (Sunburn, Canopy, DAPG, and DAPR) and fruit quality categories (FW, Brix, and pH).

Three ILs contained introgressions on two or more chromosomes: LA3958 (chromosomes 9 & 11), LA3960 (chromosomes 9, 10, & 12), and LA3965 (chromosomes 2, 10, & 11) (see Figure 3). These ILs are listed to the bottom right of Figure 2 and their corresponding chromosomal locations have been colored. Traits associated with these ILs (listed in Figure 3) were not able to be mapped to a specific chromosome. Due to overlaps between some of the introgressions in these three ILs and other ILs, there is some evidence that certain traits are associated within the regions containing these overlaps. Five traits, DAPR161, Canopy17, RYLD17, HI182, and RYLD192F, were identified in both LA3958 and LA3960. TGRAs for these traits may be located where the ILs share an introgression on chromosome 9, between marker CT112 and the telomere. Three traits, Sunburn16, Brix172R, and DAPR181, were both identified in LA3960 and LA3965. These traits also have overlapping confidence intervals, which lends evidence that they may be associated with a genomic region shared by LA3960 and LA3965 on chromosome 10 between markers TG408 and TG241.

8. Discussion

8.1 Identification of TGRAs in ILs

Two-hundred sixty-nine TGRAs were discovered in 22 tomato ILs derived from *S. habrochaites* (Figures 2 and 3). TGRAs were found for each of the 15 traits evaluated in this study. Of the TGRAs identified, the presence of wild alleles at 40 of the 268 TGRAs had a positive horticultural effect on the trait mean relative to the control E6203. However, none of the ILs contributed only positive allelic effects at TGRAs. These results indicate that the *S. habrochaites* introgressions contain both horticulturally desirable and undesirable alleles. This outcome was not unexpected as crop wild relatives have been shown to contain both horticulturally acceptable and unacceptable alleles (Fehr 1987; Grandillo and Tanksley 1996; Tanksley and Nelson 1996; Zamir 2001). Based on the results of this study it cannot be determined if undesirable horticultural traits from *S. habrochaites* are the result of close repulsion phase linkages and/or pleiotropy.

The use of any of these ILs in breeding directly would likely contribute negative effects on horticultural traits due to linkage drag. Linkage drag has been identified as a potential issue in using introgressions from *S. habrochaites* in breeding (Haggard et al., 2013; Lounsbery et al., 2016). Within the set of ILs that contain a single introgression, TGRAs with horticulturally undesirable allelic effects of the wild alleles were found on each chromosome, except chromosome 9.

Identified TGRAs may involve tight linkage, pleiotropy, or both. Due to the generally large genetic size of the introgressions present in the ILs, we cannot determine if identified

TGRAs involve either or both of these phenomena since precise TGRA locations within the introgression are unknown. For these ILs to be useful in breeding, determining if TGRAs are tightly linked loci in repulsion or are a result of pleiotropy is needed (Chen and Lubberstedt 2010). Linkage drag is commonly observed in using wild species for crop improvement (Fehr 1987; Sharma et al., 2013; Hubner and Kantar 2021). Further studies using high resolution mapping could be helpful to determine if identified TGRAs involve tight linkage among QTL (Collard et al., 2005; Lin et al., 2014). In addition, further studies could help determine if any of the TGRAs are due to pleiotropy, which in turn affects effectiveness of selection by breeders (Mackay et al., 2009; Chen and Lubberstdedt 2010). Determining if traits are controlled by tightly linked loci in repulsion or are a result of pleiotropy would be helpful for employing tomato wild relatives in breeding cultivated tomato (Lin et al., 2014).

A previous study in our lab using near-isogenic lines derived from *S. habrochaites* (Lounsbery et al., 2016) reported QTL associated with water-stress tolerance-related traits on the short arm of chromosome 9. Interestingly in our present study, the ILs with chromosome 9 introgressions did not exhibit any significant TGRAs. It is possible that the use of a less severe deficit irrigation water treatment in our study compared to Lounsbery et al. (2016) could account for the differing results. In addition, the plant material used in Lounsbery et al. (2016) was derived from a different but closely related *S. habrochaites* accession, LA1778, and the introgressions were also in a different genetic background, *S. lycopersicum* fresh market cv. Bloom T-5. These differences may explain the lack of detection of TGRAs on chromosome 9 in the ILs used in this study.

The chromosomal locations of TGRAs with ILs containing two or more introgressions could not be determined conclusively. Without overlapping coverage of a region by at least two ILs with differing introgressions, it is not possible to determine which chromosomal introgression is associated with a given trait for an IL containing more than one introgression. To resolve this issue, single introgression ILs overlapping in the same regions as the ILs containing more than one introgression would be required or through the use of genetic mapping (Eshed and Zamir 1994(b); Collard et al., 2005; Chetelat et al., 2019).

TGRAs with positive horticultural allelic effects relative to E6203 were found for traits in the fruit quality and horticultural categories. However, the introgressions the traits mapped to also contain TGRAs with negative horticultural allelic effects. Determination of a more precise location of the QTL controlling these TGRAs with a positive horticultural allelic affect would require using QTL mapping techniques (Collard et al., 2005; McCouch and Doerge, 1995; Paterson et al., 1988). Genome wide association studies could also be used to map QTL (Begum et al., 2015; Crowell et al., 2016).

None of the ILs in our study performed significantly better than E6203 for yield. For several ILs, no TGRAs associated with yield were detected, which suggests that the allelic effect of the introgression did not provide any additional benefit beyond that already present in E6203. This contrasts with the results of Lounsbery et al. (2016) who found that QTL alleles from *S. habrochaites* increased water-stress tolerance. In that prior study, a more severe deficit water treatment of 33% ETc was used and applied starting earlier in the growing season, compared to the treatment used in our study. If a future investigation was conducted with this

same set of 24 ILs, increasing the severity of the deficit water treatment may reveal potential beneficial effects of some of the wild alleles for water-stress tolerance-related traits.

A less severe deficit water treatment than that used in Lounsbery et al. (2016) was selected for this experiment because it may be feasible for processing tomato growers to potentially adopt in order to save water while maintaining yields. Implementing a moderate deficit irrigation scheme post-fruit set at mid- to late-season is more likely to be used by growers than a season-long severe deficit irrigation scheme.

Certain introgressions from *S. habrochaites* were associated with TGRAs with negative allelic effects for decreased yield. Similar findings were reported by Bernacchi et al. (1998a) in which alleles associated with *S. habrochaites* were found to significantly decrease yield. Significant reductions in yield were also associated with wild alleles from *S. pimpinellifolium* and *S. pennellii* (Tanksley et al., 1996; Eshed and Zamir, 1995). If the loci controlling the two traits possess sufficient genetic distance, backcrossing can be used as a method to recover the positive traits from the recurrent parent once linkage is broken (Fehr, 1987).

Tomato wild relatives are closely related to each other and are intercrossable (Chetelat and Ji 2007; Moyle 2008). Tomato and its wild relatives also share a high degree of synteny (Frary et al., 2016). QTL identified in syntenic regions of one wild tomato relative may suggest that another species may contain alleles effecting the same trait in the same region. These alleles in syntenic regions of other wild tomato relatives may prove useful as a resource to improve cultivated tomato.

8.2 TGRA Stability and G x E interactions / Environmental Effects

TGRA x Environment interactions are a type of G x E interaction, which can be described as the inconsistent detection of TGRAs across environments (i.e., years and locations). G x E is a result of the interaction between the expression of a plant trait phenotype and the environment it was evaluated in (Bernardo 2008; Mackay et al., 2009). The presence and magnitude of G x E interactions depend on the combination of the trait, the genetic materials, and the environment in which the plants are grown (Fehr 1987; Bernardo 2008). TGRAs may contain one or more QTL for a given trait, given the generally large genetic size of the introgressions in the ILs (Monforte and Tanksley, 2000).

TGRA x Environment interactions were detected for all traits. Both rank changes and changes in magnitude were identified among the genotypes between environments (Supplemental Table 3; Supplemental Table 4). Significant interactions between genotype and water treatment indicated that genotypes performed differently under the reduced water treatment. G x Water interactions were mostly observed in fruit quality and yield related traits. Within a TGRA, G x E did not change the direction of the horticultural desirable effect for any of the traits. Many of the trait means of the ILs were variable from environment to environment. This variability was found for each of the ILs containing TGRAs. Trait means for the fruit quality trait category tended to be the most variable. HI and the maturity traits tended to show the most consistent trait means and were detected in all the environments. The trait mean variability across environments suggests that many of the TGRAs lack environmental stability.

Variation in detection of FW QTL in differing environments has been observed in introgression lines derived from tomato wild relatives *S. galapagense* (Paterson et al., 1991) and *S. habrochaites* (Bernacchi et al., 1998(a)). FW QTL associated with the *S. galapagense*

allele were detected on chromosomes 1, 2, 3, 4, 6, 7, 9, 11, and 12, although none of the detected QTL were present in all three locations trialed (Paterson et al., 1991). In our study, FW TGRAs were identified in syntenic regions in the *S. habrochaites* ILs on chromosomes 2 (LA3922), 3 (LA3927), 4 (LA3930), 7 (LA3948 and LA3951), and 12 (LA3969). Similarly, FW QTL associated with *S. habrochaites* alleles from the same accession used in this study were identified on chromosomes 2, 3, and 4, although only in one of the three environments tested (Bernacchi et al., 1998(a)). In our study, FW TGRAs were found in the same regions on chromosomes 2 (LA3922) and 4 (LA3933) as were reported in Bernacchi et al. (1998(a)).

E6203, the recurrent parent of the IL population, is not known to be tolerant to water-stress. For the traits in which significant genotype x treatment interactions were determined, E6203 did not display consistent horticulturally superior performance under the full water treatment (Supplemental Table 2). This indicates that the level of water-stress applied in this experiment is not severe enough to adversely impact E6203. For each trait in which genotype x treatment main effect interactions were detected, E6203 exhibited little difference between genotype means between the full and reduced water treatments.

8.3 ILs for Use in Breeding

Most of the 24 ILs contained one or more TGRAs with a negative horticultural effect of the wild alleles. However, some ILs with fewer TGRAs associated with negative horticultural effects from wild alleles may be useful for breeding, especially if positive effects of the wild alleles are also present. A selection index could be created to weigh traits by their relative importance to help rank the ILs for potential utility in breeding (Luby and Shaw 2009) . Using an

index may indicate that some of the ILs may be potentially beneficial for improving certain target traits in cultivated tomato.

In general, most identified TGRAs for maturity (DAPR, DAPG) were found to contribute to later maturity. TGRAs associated with earlier maturity, which is horticulturally beneficial, were identified in LA3975 in 2016 location 1 for DAPR and LA3955 in 2017 at both locations for DAPG and DAPR. However, the introgression in LA3975 is also associated with a decrease in RYLD, which when combined with the inconsistent detection of DAPR TGRAs would make LA3975 unsuitable for use in breeding for earlier maturity. Unlike LA3975, LA3955 is not associated with a decrease in yield and contains only one other TGRA associated with decreased Brix in 2017 Location 2 under full water. LA3955 could be used to breed for earlier maturity if the stability of the TGRA is demonstrated in additional environments. Alleles from *S. habrochaites* have previously been associated with later maturity and a higher percentage of green fruit at harvest (Haggard et al., 2013; Bernacchi et al., 1998(a)). In contrast to *S. habrochaites,* QTL from *S. pimpinellifolium* on chromosomes 2, 8, and 9 are associated with earlier maturity (Grandillo and Tanksley, 1996). *S. pimpenillifolium* is likely a better candidate for use in the breeding of earlier maturity in cultivated tomato. Since the majority of tomato cultivar breeding is done in the private seed sector, it is unknown if *S. pimpinellifolium* has been used to improve maturity. To the best of our knowledge there have been no published reports on using *S. pimpinellifolium* to improve maturity in commercial tomato.

Introgressions from *S. habrochaites* were often associated with increased canopy cover and sunburn resistance. These traits were correlated with each other because in larger canopies, more leaves cover the fruit, preventing sunburn. Fruit sunburn is tissue damage from

excessive exposure to high intensity sunlight. Fruit that are not shaded by plant vegetation are at higher risk of developing sunburn (Racsko and Schrader, 2012). ILs containing TGRAs with increased sunburn resistance as well as increased canopy cover also contained TGRAs for decreased yield. This suggests there is a biological tradeoff between vegetative biomass production and fruit production. Some ILs containing TGRAs for increased sunburn resistance and increased canopy cover also had increased BioM (LA3921 and LA3948). Similar results were found in *S. pennelii* ILs in which increased plant weight was generally associated with a decrease in yield (Eshed and Zamir, 1995).

TGRAs were found for increased Brix on chromosomes 1, 2, 5, 7, 8, 9, 10, 11, and 12 (Figure 2). All of the ILs containing positive TGRAs for Brix also contained negative TGRAs for one or more other traits. In addition, the presence of TGRAs for Brix differed from environment to environment. Only one IL, LA3913 (chromosome 1) contained a TGRA for Brix in most environments. Brix QTL from *S. habrochaites* acc. LA 1777 were previously mapped on chromosomes 3, 4, 5, 6, and 9 (Monforte et al., 2001; Bernacchi et al., 1998(a)). Additionally, Bernacchi et al. (1998(b)) found Brix QTL from *S. habrochaites* on chromosomes 3 and 5. These results further illustrate the environmental and genotype x environment effects on Brix.

Other tomato wild relatives have been found to contain QTL for increased Brix. The large number of identified QTL from multiple tomato wild relatives associated with increased Brix illustrates that there are various genetic resources for the improvement of Brix in cultivated tomato. Introgressions from *S. galapagense* contain QTL for increased Brix on chromosomes 2, 3, 6, 7, and 9 (Paterson et al., 1991). The QTL from *S. galapaganse* are syntenic with introgressions from *S. habrochaites* on chromosomes 2 (LA3922), 7 (LA3948), and 9

(LA3956 and LA3957). Of these ILs that share syntenic regions with a region from *S. galapaganse* associated with increased Brix content, LA3948 contained a TGRA for increased Brix. *S. pennellii* has also been found to contribute QTL associated with increased Brix on each tomato chromosome (Eshed and Zamir, 1995). In addition to the QTL mapped by Eshed and Zamir (1995), QTL for increased Brix were also mapped to chromosomes 7, 8 and 12 in *S. pennellii* (Ikeda et al., 2013; Sacco et al., 2013) and were identified in regions syntenic to LA3951 (chromosome 7), LA3953 (chr 8), LA3955 (chr 8) and LA3969 (chr 12). LA3953 and LA3955 had TGRAs associated with Brix. In contrast to what was found in Ikeda et al. (2013), LA3953 showed a decrease in Brix while the larger introgression in LA3955 had a TGRAs associated with an increase in Brix. QTL for increased Brix have also been detected on chromosome 4 in *S. pennellii* and *S. peruvianum* (Monforte et al., 2001). None of the ILs in our study contained an introgression in a region syntenic to *S. pennellii* and *S. peruvianum* regions in which QTL were detected. QTL associated with increased Brix have been found in introgressions on chromosomes 3, 6, and 9 from *S. pimpinellifolium* (Grandillo and Tanksley, 1996).

The ILs were not significantly different from E6203 for carotenoids (CRT) except LA3921, in which it was significantly lower than E6203. The fruit of this IL had a yellow-orange color instead of red, likely due to the decreased concentration of lycopene, the pigment that gives tomato its red color (Rao and Rao, 2007). Because none of the ILs had a higher CRT content than E6203, they cannot be used in breeding to increase the CRT concentration present in the fruit.

Introgressions would likely have different effects on trait expression and genotype performance in different genetic backgrounds (Fehr 1987; Eshed et al., 1996; Tanskley et al., 1996). E6203, the recurrent parent of the ILs, is only one of many possible genetic backgrounds in cultivated tomato. Introgressions from wild species can show variable effects on trait performance, depending on the genetic background to which they are transferred into (Nyine et al., 2021). Once a TGRA or QTL is identified, further testing in other genetic backgrounds is essential to assess the stability of QTL expression (Eshed et al., 1996; Tanksley et al., 1996). Interactions between the TGRA or QTL and the genetic background they are located in can have effects on trait expression and performance of genotypes (Eshed et al., 1996). Further testing of introgressions transferred into other genetic backgrounds would be helpful to determine which *S. habrochaites* introgressions would be beneficial for improvement of tomato. Backcrossing ILs with horticulturally desirable traits into elite breeding material may be a method to transfer selected traits from *S. habrochaites* into the processing tomato germplasm.

9. Acknowledgements

This research was supported by the US Department of Agriculture National Institute of Food and Agriculture competitive grants. We thank the St. Clair lab members for contributions to field and lab data collection. We thank Drs. Li Tian and Jingjing Liu for their expertise in development of the carotenoids assay. We also thank the UC Davis Plant Sciences field facility crew for assistance with field experiments.

10. Author Contributions

Bryce A. Kubond: conceptualization, conduct experiments, formal analysis, investigation, project administration, writing- original draft, writing- review editing. Dina A. St. Clair: conceptualization, funding acquisition, project administration, writing- review editing.

11. Conflict of Interest Statement

There are no known conflicts of interest.

12. Supplemental Material

Supplemental Table S1:

Genotype x Environment interactions detected in traits evaluated in a set of 24 ILs derived from *S. habrochaites* grown in the field under two water treatments. Genotype x Location and Genotype x Treatment interactions are listed for each trait they were identified with. If a Genotype x Location or Genotype x Treatment interaction was detected for a particular trait, the trait is listed next to the year(s) and or locations the interaction with genotype was detected in.

Supplemental Table S2:

Trait means and Tukey mean separation groups for 24 ILs derived from *S. habrochaites.* ILs and control E6203 were tested in the field under two water treatments. IL accessions are listed to the left of the table. The chromosome(s) where the introgression(s) for each accession is located are listed to the right of the accession. Under each trait, trait means for each accession

are listed to the left. Tukey groups are listed to the right of each trait mean. The Coefficient of Variation (CV) for each trait is listed at the bottom of the table beneath each trait.

Supplemental Table S3:

Spearman rank correlations were computed for each pairwise combination of categorical traits (Canopy, Habit, and Sunburn) in the ILs plus E6203. Blue shaded cells indicate a horticulturally desirable correlation. Orange shaded cells indicate a horticulturally undesirable correlation. Six levels of shading for each color were used, with the range spanning the lightest encompassing $0.40 < |r| < 0.49$ to the darkest encompassing $|r| > 0.90$. Levels of shading are in increments of 0.1. Gray cells indicate non-significant correlations.

Supplemental Table S4:

Detected TGRAs in 24 ILs derived from *S. habrochaites* and evaluated in the field under two water treatments in three years. The identified TGRAs are listed with the IL accession that the TGRA is associated with, the chromosome(s) containing the introgression(s), the flanking markers of the introgression(s), the *P* value, and the percent difference in trait value from the recurrent parent control E6203.

13. References

Anthon, G.E., LeStrange, M. and Barrett D.M. 2011. Changes in pH, acids, sugars and other quality parameter during extended vine holding of ripe processing tomatoes. *Journal of the Science of Food and Agriculture*. 91: 1175-1181. https://doi.org/10.1002/jsfa.4312

Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2015. High-resolution mapping of a major effect QTL from wild tomato *Solanum habrochaites* that influences water relations under root chilling. *Theoretical and Applied Genetics*. 128: 1713-1724. https://doi.org/10.1007/s00122-015-2540-y

Bai, Y., Kissoudis, C., Yan, Z., Visser, R.G.F. and Van der Linden, G. 2018. Plant behaviour under combined stress: tomato responses to combined salinity and pathogen stress. *The Plant Journal*. 93: 781-793. https://doi.org/10.1111/tpj.13800

Barrett, D.M., Weakley, C., Diaz, J.V. and Watnik, M. 2007. Qualitative and nutritional differences in processing tomatoes grown under commercial organic and conventional production systems. *Journal of Food Science*. 72(9): 441-451. https://doi.org/10.1111/j.1750-3841.2007.00500.x

Bates, D., Maechler, M., Bolker, B., and Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1): 1-48. https://doi.org/10.48550/arXiv.1406.5823

Bernacchi, D., Beck-Bunn, T., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley S.D. 1998(a). Advanced backcross QTL analysis in tomato. I. identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theoretical and Applied Genetics*. 97: 381-397. https://doi.org/10.1007/s001220050908

Bernacchi, D., Beck-Bunn, T., Emmatty, D., Eshed, Y., Inai, S., Lopez, J., Petiard, V., Sayama, H., Uhlig, J., Zamir, D. and Tanksley S.D. 1998(b). Advanced backcross QTL analysis of tomato II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *Theoretical and Applied Genetics*. 97: 170-180. https://doi.org/10.1007/s001220050882

Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: learning form the last 20 years. *Crop Science*. 48(5): 1649-1664. https://doi.org/10.2135/cropsci2008.03.0131

Begum, H., Spindel, J.E., Lalusin, A., Borromeo, T., Gregorio, G., Hernandez, J., Virk, P., Collard, B. and McCouch, S.R. 2015. Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice (*Oryza sativa*). *PLOS One*. 10(3): 1-19. https://doi.org/10.1371/journal.pone.0119873

Canady, M.A., Meglic, V. and Chetelat, R.T. 2005. A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome*. 48: 685-697. https://doi.org/10.1139/g05-032

Chen, Y. and Lubberstedt, T. 2010. Molecular basis of trait correlations. *Trends in Plant Science*. 15: 454-461. https://doi.org/10.1016/j.tplants.2010.05.004

Chetelat, R.T., and Ji, Y. 2007. Cytogenetics and evolution. In *Genetic Improvement of Solanaceous Crops* (Razdan, M.K. and Mattoo, A.K., eds). Science Publishers. Enfield, NH. pp 77-112.

Chetelat, R.T., Qin, X., Tan, M., Burkart-Waco, D., Moritama, Y., Huo, X., Wills, T. and Pertuzé R. 2019. Introgression lines of *Solanum sitiens*, a wild nightshade of the Atacama Desert, in the genome of cultivated tomato. *The Plant Journal*. 100(4): 836-850. https://doi.org/10.1111/tpj.14460

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B. and Pang E.C.K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*. 142: 169-196. https://doi.org/10.1007/s10681-005-1681-5

Crowell, S., Korniliev, P., Falcao, A., Ismail, A., Gregorio, G., Mezey, J. and McCouch, S. 2016. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nature Communications.* 7:10527. https://doi.org/10.1038/ncomms10527

Diffenbaugh, N.S., Swain, D.L. and Touma D. 2015. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*. 112(13): 3931-3936. https://doi.org/10.1073/pnas.1422385112

Eshed, Y. and Zamir, D. 1994 (a). A genomic library of *Lycopersicon pennellii* in *L. esculentum*: a tool for fine mapping of genes. *Euphytica*, 79: 175-180. https://doi.org/10.1007/BF00022516

Eshed, Y., and Zamir, D. 1994 (b). Introgressions from *Lycopersicon pennellii* can improve the soluble-solids yield of tomato hybrids. *Theoretical and Applied Genetics*. *88*(6-7): 891-897. https://doi.org/10.1007/BF01254002

Eshed, Y., and Zamir, D. 1995. An Introgression Line Population of *Lycopersicon pennelii* in the Cultivated Tomato Enables the Identification and Fine Mapping of Yield-Associated QTL. *Genetics*. 141: 1147-1162. https://doi.org/10.1093/genetics/141.3.1147

Eshed, Y., Gera, G. and Zamir, D. 1996. A genome-wide search for wild-species alleles that increase horticultural yield of processing tomatoes. *Theoretical and Applied Genetics*. 93: 877-886. https://doi.org/10.1007/BF00224089

Fehr, W.R. 1987. *Principles of Cultivar Development, vol. 1: Theory and Technique*. Iowa State University Press. Ames, Iowa.

Food and Agriculture Organization of the United Nations. 2021. *FAOSTAT Crops*. FAOSTAT, Rome, Italy. http://www.fao.org/faostat/en/#data/QC (accessed 22 June 2021).

Frary, A., Doganlar, S., and Frary A. 2016. Synteny Among Solanaceae Genomes. In: Causse, M., Giovanni, J., Bouzayen, M., and Zouine, M., editors, Compendium of Plant Genomes. Springer-Verlag, Berlin, Heidelberger Platz. p. 217-243. https://doi.org/10.1007/978-3-662-53389-5_12

Fulton, T.M., Beck-Bunn, T., Emmatty, D., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley, S.D. 1997. QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the

cultivated tomato and comparisons with QTLs found in other wild species. *Theoretical and Applied Genetics*. 95: 881-894. https://doi.org/10.1007/s001220050639

Fulton, T.M., Grandillo, S., Beck-Bunn, T., Fridman, E., Frampton, A., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley, S.D. 2000. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics*. 100: 1025-1042. https://doi.org/10.1007/s001220051384

Glas, J.J., Schimmel, B.C.J., Alba, J.M., Escobar-Bravo, R., Schurrink, R.C. and Kant, M.R. 2012. Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*. 13(12): 17077-17103. https://doi.org/10.3390/ijms131217077

Goodstal F.J., Kohler G.R., Randall L.B., Bloom A.J. and St. Clair D.A. 2005. A major QTL introgressed from wild *Lycopersicon hirsutum* confers chilling tolerance to cultivated tomato (*Lycopersicon esculentum*). *Theoretical and Applied Genetics*, 111: 898-905. https://doi.org/10.1007/s00122-005-0015-2

Grandillo, S. and Tanksley, S.D. 1996. QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersion pimpinellifolium*. *Theoretical and Applied Genetics*. 92: 935-951. https://doi.org/10.1007/BF00224033

Groh, A.M., Kubond, B.A., and St. Clair, D.A. 2022. Fine mapping of QTL for water use efficiency-related traits on chromosome 9 of *Solanum habrochaites* in the field. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20828

Haggard, J.E., Johnson, E.B., and St. Clair D.A. 2013. Linkage relationships among multiple QTL for horticultural traits and Late Blight (*P. infestans*) resistance on chromosome 5 introgressed from wild tomato *Solanum habrochaites*. *G3.* 3(12): 2131-2146. https://doi.org/10.1534/g3.113.007195

Hartman, J.B. and St. Clair, D.A. 1999. Combining ability for beet armyworm, *Spodoptera exigua*, resistance and horticultural traits of selected *Lycopersicon pennellii*-derived inbred backcross lines of tomato. *Plant Breeding*, 118(6): 523-530. https://doi.org/10.1046/j.1439-0523.1999.00437.x

Hartz, T.K., Miyao, G., Mickler., Lestrange, M., Stoddard., Nuñez, J., and Aegerter B. 2008. Processing Tomato Production in California. *University of California Agriculture and Natural Resources*, doi: 10.3733/ucanr.7228

Hubner, S. and Kantar, M.B. 2021. Tapping diversity from the wild: from sampling to implementation. *Frontiers in Plant Science*. 12: 626565. https://doi.org/10.3389/fpls.2021.626565

Ikeda, H., Hiraga, M., Shirasawa, K., Nishiyama, M., Kanahama, K. and Kanayama Y. 2013. Analysis of a tomato introgression line, IL8-3, with increased Brix content. *Scientia Horticulturae*. 153: 103-108. https://doi.org/10.1016/j.scienta.2013.02.006

Laur, L. M., and Tian, L. 2011. Provitamin A and vitamin C contents in selected California-grown cantaloupe and honeydew melons and imported melons. *Journal of Food Composition and Analysis*. *24*(2): 194-201. https://doi.org/10.1016/j.jfca.2010.07.009

Lee, J., De Gryze, S., Six, J. 2010. Effect of climate change on field crop production in California's Central Valley. *Climate Change*. 109: 335-353. https://doi.org/10.1007/s10584-011-0305-4

Lenth, R.V. 2016. Least-Squares Means: The R package lsmeans. *Journal of Statistical Software*, 69(1): 1-33. https://doi.org/10.18637/jss.v069.i01

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., Wang, X., Huang, Z., Li, J., Zhang, C., Wang, T., Zhang, Y., Wang, A., Zhang, Y., Lin, K., Li, C., Xiong, G., Xue, Y., Mazzucato, A., Causse, M., Fei, Z., Giovanni, J.J., Chetelat, R.T., Zamir, D., Stadler, T., Zhibiao, Y., Du, Y. and Huang, S. 2014. Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics*. 46(11): 1220-1226. https://doi.org/10.1038/ng.3117

Lounsbery, J.K., Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2016. Quantitative Trait Loci for water-stress tolerance traits localize on chromosome 9 of wild tomato. *Crop Science*, 56: 1514-1525. https://doi.org/10.2135/cropsci2015.07.0432

Luby, J.J., and Shaw, D.V. 2009. Plant breeders' perspectives on improving yield and quality traits in horticultural food crops. *HortScience*. 44(1): 20-22. https://doi.org/10.21273/HORTSCI.44.1.20

Mackay, T.F.C., Stone, E.A. and Ayroles, J.F. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews, Genetics*. 10: 565-577. https://doi.org/10.1038/nrg2612

Mahoney, K., Scott, J.D., Alexander, M., McCray, R., Hughes, M., Swales, D. and Bukovsky, M. 2021. Cool season precipitation projections for California and the Western United States in NA-CORDEX models. *Climate Dynamics*. 56: 3081-3102. https://doi.org/10.1007/s00382-021-05632-z

McCouch, S.R., and Doerge, R.W. 1995. QTL mapping in rice. *Trends Genetics*. 11: 482-487. https://doi.org/10.1016/S0168-9525(00)89157-X

Miller, J.C. and Tanksley, S.D. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics*. 80: 437-448. https://doi.org/10.1007/BF00226743

Monforte, A.J., and Tanksley, S.D. 2000. Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicum hirsutum* genome in a *L.*

esculentum genetic background: a tool for gene mapping and gene discovery. *Genome*, 43(5): 803-813. https://doi.org/10.1139/g00-043

Monforte, A.J., Friedman, E., Zamir, D. and Tanksley, S.D. 2001. Comparison of a set of allelic QTL-NILs for chromosome 4 of tomato: Deductions about natural variation and implications for germplasm utilization. *Theoretical and Applied Genetics*. 102:572-590. https://doi.org/10.1007/s001220051684

Mote, P.W., Hamlet, A.F., Clark, M.P. and Lettenmaier, D.P. 2005. Declining mountain snowpack in western North America. *Bulletin of the American Meterological Society*. 86(1):39-49. https://doi.org/10.1175/BAMS-86-1-39

Moyle, L.C. 2008. Ecological and evolutionary genomics in the wild tomatoes (Solanum sect. Lycopersicon). *Evolution*. 62(12): 2995-3013. https://doi.org/10.1111/j.1558-5646.2008.00487.x

National Agricultural Statistics Service. 2021. *Statistics by subject*. USDA, Washington DC. http://www.nass.usda.gov/Statistics_by_Subject/?sector=CROPS (accessed 29 Mar. 2021).

Nyine, M., Adhikari, E., Clinesmith, M., Aiken, R., Betzen, B., Wang, W., Davidson, D., Yu, Z., Guo, Y., He, F., Akhunova, A., Jordan, K.W., Fritz, A.K. and Ahkunov, E. 2021. The haplotype-based analysis of *Aegilops tauschii* introgressions into hard red winter wheat and its impact on productivity traits. *Frontiers in Plant Science*. 12: 716955. https://doi.org/10.3389/fpls.2021.716955

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln S.E. and Tanksley, S.D. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature*. 335: 721-726. https://doi.org/10.1038/335721a0

Paterson, A.H., Damon, S., Hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., Lander, E.S. and Tanksley, S.D. 1991. Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations, and environments. *Genetics*. 127: 181-197. https://doi.org/10.1093/genetics/127.1.181

Payero, J.O., Tarkalson, D.D., Irmak, S., Davison, D. and Petersen, J.L. 2009. Effect of timing of a deficit-irrigation allocation on corn evapotranspiration, yield, water use efficiency and dry mass. *Agricultural Water Management*. 96: 1387-1397. https://doi.org/10.1016/j.agwat.2009.03.022

Tanksley, S.D., and Nelson, J.C. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics*. 92: 191-203. https://doi.org/10.1007/BF00223376

Tanksley, S.D., Grandillo, S., Fulton, T.M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. and Beck-Bunn T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of

tomato and its wild relative *S. pimpinellifolium*. *Theoretical and Applied Genetics*. 92: 213-224. https://doi.org/10.1007/BF00223378

R Core Team 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Racsko, J. and Schrader, L.E. 2012. Sunburn of apple fruit: historical background, recent advances and future perspectives. *Critical Reviews in Plant Sciences*. 31(6): 455-504. https://doi.org/10.1080/07352689.2012.696453

Rao, A.V. and Rao, L.G. 2007. Carotenoids and human health. *Pharmacological Research*. 55(3): 207-216. https://doi.org/10.1016/j.phrs.2007.01.012

Richards, R.A. 2006. Physiological traits used in the breeding of new cultivars for water-scarce environments. *Agricultural Water Management*. 80: 197-211. https://doi.org/10.1016/j.agwat.2005.07.013

Rick C.M. 1973. Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. In A.M. Srb (Ed.). *Genes, Enzymes, and Populations*. p. 255-269. Boston, MA, USA: Springer. https://doi.org/10.1007/978-1-4684-2880-3_17

Rick C.M. 1983. Genetic variability in tomato species. *Plant Mol. Biol. Rep.* 1: 81-87. https://doi.org/10.1007/BF02680303

Ruan, Y.L., Patrick, J.W., Bouzayen, M., Osorio, S. and Fernie, A.R. 2012. Molecular regulation of seed and fruit set. *Trends in Plant Science.* 17(11): 656-665. https://doi.org/10.1016/j.tplants.2012.06.005

Sacco, A., Di Matteo, A., Lombardi, N., Trotta, N., Punzo, B., Mari, A. and Barone, A. 2013. Quantitative trait loci pyramiding for fruit quality traits in tomato. *Molecular Breeding*. 31: 217-222. https://doi.org/10.1007/s11032-012-9763-2

Schouten, H.J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y. and Visser, R.G.F. 2019. Breeding has increased the diversity of cultivated tomato in The Netherlands. *Frontiers in Plant Science*. 10: 1606. https://doi.org/10.3389/fpls.2019.01606

Sharma S., Upadhyaya, H.D., Varshney, R.K. and Gowda, C.L.L. 2013. Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Frontiers in Plant Science*. 4: 309. https://doi.org/10.3389/fpls.2013.00309

Spooner, D.M., Peralta, E. and Knapp, S. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes (*Solanum L.* section *Lycopersicon* (Mill.) Wettst.). *Taxon*. 54:43-61.47. https://doi.org/10.2307/25065301

Truco, M.J., Randall, L.B., Bloom A.J. and St. Clair D.A. 2000. Detection of QTLs associated with shoot wilting and root ammonium uptake under chilling temperatures in an interspecific

backcross population from *Lycopersicum esculentum* x *L. hirsutum*. *Theoretical and Applied Genetics*. 101:1092-1092. https://doi.org/10.1007/s001220051583

Tuberosa, R. 2012. Phenotyping for drought tolerance of crops in the genomic era. *Frontiers in Physiology*. 3(347): 1-26. https://doi.org/10.3389/fphys.2012.00347

Yang, X.H., Caro, M., Hutton, S.F., Scott, J.W., Guo, Y.M., Wang, X.X., Rashid, M.H., Szinay, D., de Jong, H., Visser, R.G.F., Bai, Y.L. and Du, Y.C. 2014. Fine mapping of the tomato yellow leaf curl virus resistance gene *Ty-2* on chromosome 11 of tomato. *Molecular Breeding.* 34(2): 749-760. https://doi.org/10.1007/s11032-014-0072-9

Zamir, D. 2001. Improving plant breeding with exotic genetic libraries. *Nature Reviews*. 2: 983-989. https://doi.org/10.1007/s11032-014-0072-9

14. Figures and Figure Legends

Figure 1:

Spearman rank correlations were computed for each pairwise combination of water-stress tolerance-related traits (highlighted in yellow) and horticultural traits (highlighted in green) for the 24 ILs plus E6203. Blue shaded cells indicate a horticulturally desirable correlation. Orange shaded cells indicate a horticulturally undesirable correlation. Six levels of shading for each color were used, with the range spanning the lightest encompassing $0.40 < |r| < 0.49$ to the darkest encompassing $|r| > 0.90$. Levels of shading are in increments of 0.1. Gray cells indicate non-significant correlations.

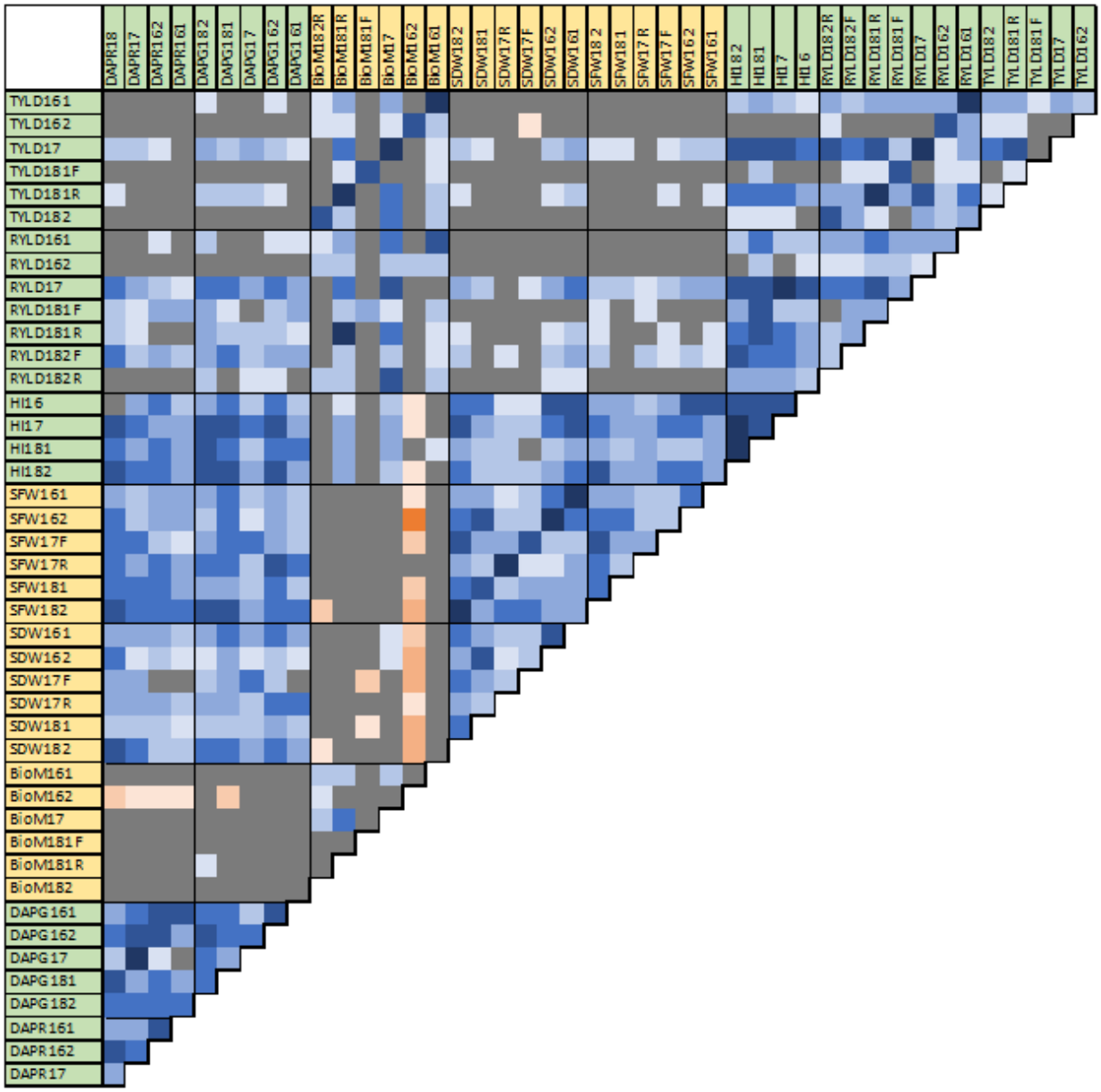

DAPR18
DAPR17
DAPR162
DAPR161
DAPG182
DAPG181
DAPG17
DAPG162
DAPG161
BioM182R
BioM181R
BioM181F
BioM17
BioM162
BioM161
SDW182
SDW181
SDW17R
SDW17F
SDW162
SDW161
SPW182
SPW181
SPW17R
SPW17F
SPW162
SPW161
Hi182
Hi181
Hi17
Hi16
RYLD182R
RYLD182F
RYLD181R
RYLD181F
RYLD17
RYLD162
RYLD161
TYLD182
TYLD181R
TYLD181F
TYLD17
TYLD162
TYLD161
TYLD181F
TYLD181R
TYLD182
RYLD161
RYLD162
RYLD17
RYLD181F
RYLD181R
RYLD182F
RYLD182R
Hi16
Hi17
Hi181
Hi182
SPW161
SPW162
SPW17F
SPW17R
SPW181
SPW182
SDW161
SDW162
SDW17F
SDW17R
SDW181
SDW182
BioM161
BioM162
BioM17
BioM181F
BioM181R
BioM182
DAPG161
DAPG162
DAPG17
DAPG181
DAPG182
DAPR161
DAPR162
DAPR17

Figure 2:

Chromosome maps of locations of significant TGRAs detected in 24 ILs from *S. habrochaites*. The 11 chromosomes represented in various introgressions in the 24 ILs included in this study are displayed, with marker positions listed to the left of each chromosome. Colored portions of the chromosome indicate the position of each introgression, the corresponding IL is listed to the right in the same color as it appears on the chromosome. Beneath, and to the right, of each IL are the traits that have a significant trait-genomic region associations (TGRAs). For traits in which each year, location, and treatment have a significant association, the year prefix has been dropped. After each trait code, the horticultural effect of the wild species alleles at the TGRA is listed within parenthesis, with a negative (-) being undesirable and a positive (+) being desirable horticulturally. Traits are listed for each of the single introgressions.

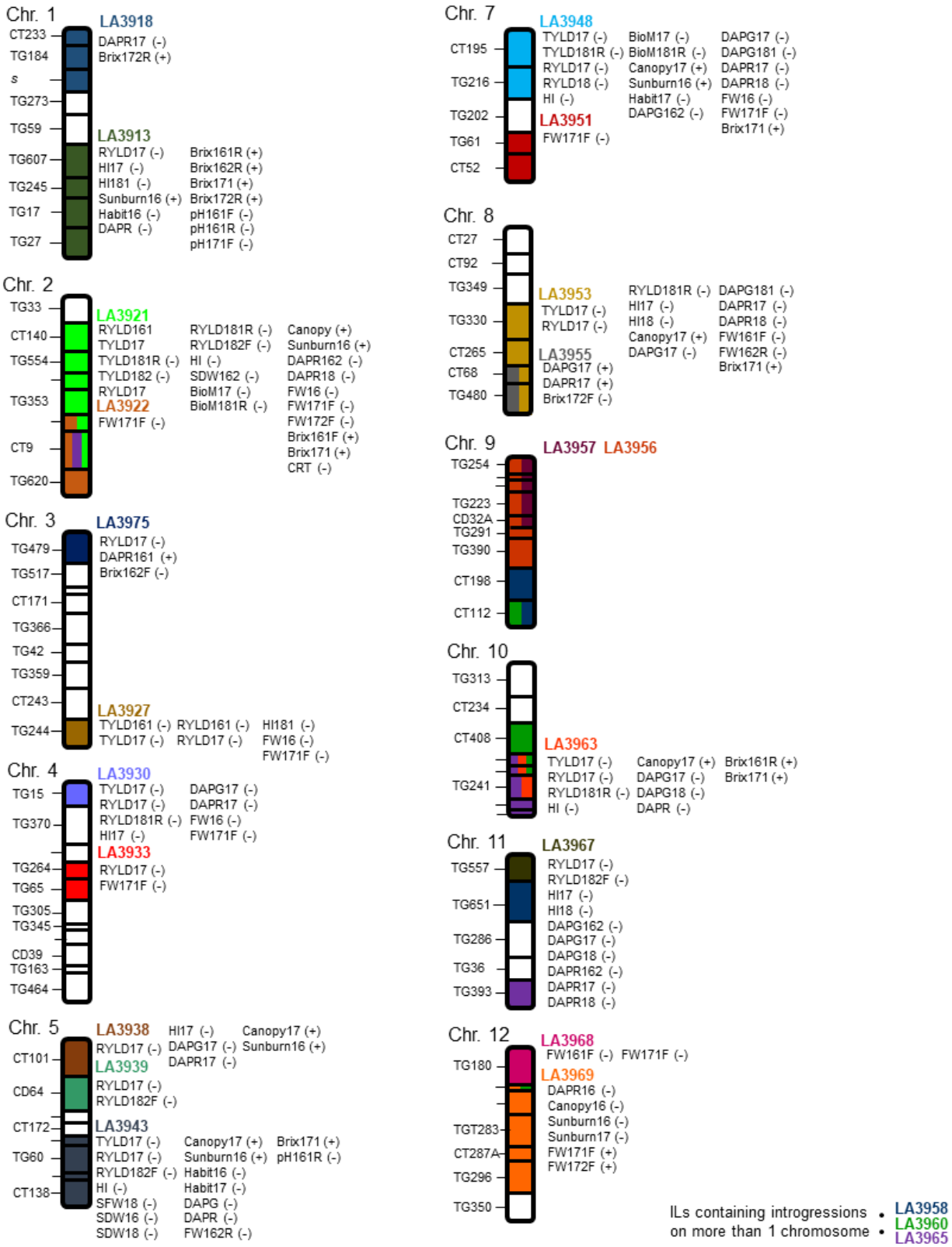
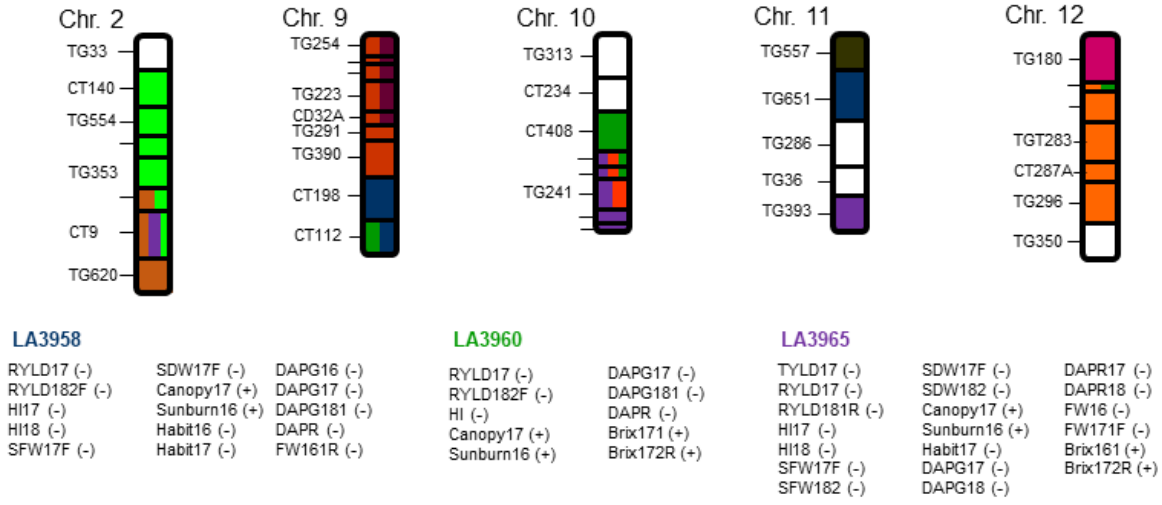

Chr. 1
CT233
TG184
s
TG273
TG59
TG607
TG245
TG17
TG27
LA3918
DAPR17 (-)
Brix172R (+)
LA3913
RYLD17 (-)
HI17 (-)
HI181 (-)
Sunburn16 (+)
Habit16 (-)
DAPR (-)
Brix161R (+)
Brix162R (+)
Brix171 (+)
Brix172R (+)
pH161F (-)
pH161R (-)
pH171F (-)
Chr. 2
TG33
CT140
TG554
TG353
CT9
TG620
LA3921
RYLD161
TYLD17
TYLD181R (-)
TYLD182 (-)
RYLD17
LA3922
FW171F (-)
RYLD181R (-)
RYLD182F (-)
HI (-)
SDW162 (-)
BioM17 (-)
BioM181R (-)
Canopy (+)
Sunburn16 (+)
DAPR162 (-)
DAPR18 (-)
FW16 (-)
FW171F (-)
FW172F (-)
Brix161F (+)
Brix171 (+)
CRT (-)
Chr. 3
TG479
TG517
CT171
TG366
TG42
TG359
CT243
TG244
LA3975
RYLD17 (-)
DAPR161 (+)
Brix162F (-)
LA3927
TYLD161 (-)
TYLD17 (-)
RYLD161 (-)
RYLD17 (-)
HI181 (-)
FW16 (-)
FW171F (-)
Chr. 4
TG15
TG370
TG264
TG65
TG305
TG345
CD39
TG163
TG464
LA3930
TYLD17 (-)
RYLD17 (-)
RYLD181R (-)
HI17 (-)
DAPG17 (-)
DAPR17 (-)
FW16 (-)
FW171F (-)
LA3933
RYLD17 (-)
FW171F (-)
Chr. 5
CT101
CD64
CT172
TG60
CT138
LA3938
RYLD17 (-)
LA3939
RYLD17 (-)
RYLD182F (-)
LA3943
TYLD17 (-)
RYLD17 (-)
RYLD182F (-)
HI (-)
SFW18 (-)
SDW16 (-)
SDW18 (-)
HI17 (-)
DAPG17 (-)
DAPR17 (-)
Canopy17 (+)
Sunburn16 (+)
Habit16 (-)
Habit17 (-)
DAPG (-)
DAPR (-)
FW162R (-)
Brix171 (+)
pH161R (-)
Chr. 7
CT195
TG216
TG202
TG61
CT52
LA3948
TYLD17 (-)
TYLD181R (-)
RYLD17 (-)
RYLD18 (-)
HI (-)
LA3951
FW171F (-)
BioM17 (-)
BioM181R (-)
Canopy17 (+)
Sunburn16 (+)
Habit17 (-)
DAPG162 (-)
DAPG17 (-)
DAPG181 (-)
DAPR17 (-)
DAPR18 (-)
FW16 (-)
FW171F (-)
Brix171 (+)
Chr. 8
CT27
CT92
TG349
TG330
CT265
CT68
TG480
LA3953
TYLD17 (-)
RYLD17 (-)
LA3955
DAPG17 (+)
DAPR17 (+)
Brix172F (-)
RYLD181R (-)
HI17 (-)
HI18 (-)
Canopy17 (+)
DAPG17 (-)
DAPG181 (-)
DAPR17 (-)
DAPR18 (-)
FW161F (-)
FW162R (-)
Brix171 (+)
Chr. 9
TG254
TG223
CD32A
TG291
TG390
CT198
CT112
LA3957
LA3956
Chr. 10
TG313
CT234
CT408
TG241
LA3963
TYLD17 (-)
RYLD17 (-)
RYLD181R (-)
HI (-)
Canopy17 (+)
DAPG17 (-)
DAPG18 (-)
DAPR (-)
Brix161R (+)
Brix171 (+)
Chr. 11
TG557
TG651
TG286
TG36
TG393
LA3967
RYLD17 (-)
RYLD182F (-)
HI17 (-)
HI18 (-)
DAPG162 (-)
DAPG17 (-)
DAPG18 (-)
DAPR162 (-)
DAPR17 (-)
DAPR18 (-)
Chr. 12
TG180
TGT283
CT287A
TG296
TG350
LA3968
FW161F (-)
FW171F (-)
LA3969
DAPR16 (-)
Canopy16 (-)
Sunburn16 (-)
Sunburn17 (-)
FW171F (+)
FW172F (+)
ILs containing introgressions on more than 1 chromosome
LA3958
LA3960
LA3965

Figure 3:

Chromosome maps displaying locations of three ILs with introgressions on two or more chromosomes. The five chromosomes representing introgressions from one of the three ILs with introgressions on two or more chromosomes are displayed with marker positions listed to the left of each chromosome. Colored portions of the chromosome indicate the position of each introgression, and the corresponding IL is listed below in the same color as it appears on the chromosome. Beneath each IL are the TGRAs that are significantly associated with it. For traits in which each year, location, and treatment have a significant association, the year prefix has been dropped. After each trait code the horticultural effect of the wild species alleles at the TGRA is listed within parenthesis, with a negative (-) being undesirable and a positive (+) being desirable horticulturally. Positions of single introgression ILs are displayed for reference.



15. Tables and Table Legends

Table 1:

Introgression lines derived from *S. habrochaites* accession LA1777 in the recurrent parent background of *S. lycopersicum* cv. E6203 used in this study. All lines were obtained from the TGRC.

Chromosomal origin of introgression:	Introgression Line accession number (TGRC):
1	LA3913, LA3918
2	LA3921, LA3922, LA3965
3	LA3975, LA3927
4	LA3930, LA3933
5	LA3938, LA3939, LA3943
7	LA3948, LA3951
8	LA3953, LA3955
9	LA3956, LA3957, LA3958, LA3960
10	LA3960, LA3963, LA3965
11	LA3958, LA3965, LA3967
12	LA3960, LA3968, LA3969

Table 2:

Traits evaluated in a set of 24 ILs grown in the field in three years (2016-2018), by category. All traits were evaluated in 2018, except those in the fruit quality category.

Trait Category	Trait Code	Description
Water-Stress Tolerance-Related	SDW	Shoot dry weight (biomass, in g)
	SFW	Shoot fresh weight (biomass, in g)

	BioM	Total fresh above ground biomass at harvest (g)
Horticultural	RYLD	Ripe fruit yield (kg)
	TYLD	Total fruit yield (kg)
	DAPG	Days after planting to first green fruit
	DAPR	Days after planting to first ripe fruit
	Canopy	Leaf canopy cover (score 1-5, sparse to dense)
	Sunburn	Degree of fruit sunburn (score 1-5, none to severe)
	Habit	Plant growth habit (score 1-5, prostrate to erect)
	HI	Harvest index (ratio of ripe fruit yield to shoot fresh weight and total fruit yield)
Fruit Quality	FW	Weight (g) of 25 fruits
	pH	pH of tomato fruit slurry
	Brix	Soluble solids content of tomato fruit slurry (degrees Brix)
	CRT	Total carotenoids content (ug/ml) of tomato fruit slurry

Table 3:

Summary of F test values for ANOVAs performed on trait data for the set of 24 ILs plus E6203. A hyphen (-) indicates not included in the model, and a blank space in the water treatment column indicates that data for water treatments were combined.

				F-Test Values		
Trait Code	Year	Field Location(s)	Water Treatment	Genotype	Location	Water Treatment
TYLD	2016	1		5.74***	-	8.09**
	2016	2		4.07***	-	3.00ns
	2017	2		9.72***	-	0.00ns
	2018	1	Full	1.65ns	-	-
	2018	1	Reduced	6.92***	-	-
	2018	2		3.29***	-	7.65***

RYLD	2016	1		6.30***	-	0.284ns
	2016	2		4.14***	-	2.89ns
	2017	2		16.65***	-	0.03ns
	2018	1	Full	2.38**	-	-
	2018	1	Reduced	8.73***	-	-
	2018	2	Full	4.34***	-	-
	2018	2	Reduced	3.57***	-	-
HI	2016	1, 2		26.30***	14.17**	5.52*
	2017	2		7.69***	-	0.00ns
	2018	1		13.36***	-	5.43*
	2018	2		14.79***	-	0.20ns
SFW	2016	1		6.79***	-	7.74**
	2016	2		5.78***	-	0.00ns
	2017	2	Full	5.15***	-	-
	2017	2	Reduced	2.30**	-	-
	2018	1		3.38***	-	0.03ns
	2018	2		9.57***	-	6.37*
SDW	2016	1		11.53***	-	3.18ns
	2016	2		7.23***	-	0.08ns
	2017	2	Full	5.20***	-	-
	2017	2	Reduced	1.89*	-	-
	2018	1		5.42***	-	4.57*
	2018	2		10.96***	-	11.34***
BioM	2016	1		2.33***	-	9.55**
	2016	2		2.74***	-	1.94ns
	2017	2		5.37***	-	0.00ns
	2018	1	Full	1.34ns	-	-
	2018	1	Reduced	5.14***	-	-
	2018	2		2.79***	-	8.57**
Canopy	2016	1, 2		23.79***	0.40ns	0.26ns
	2017	1, 2		23.60***	0.30ns	1.05ns
Sunburn	2016	1, 2		20.03***	1.14ns	0.61ns
	2017	1, 2		13.83***	1.34ns	0.26ns
Habit	2016	1, 2		31.46***	7.57ns	0.72ns
	2017	1		23.25***	-	9.82**
	2017	2		18.54***	-	2.21ns
DAPG	2016	1		12.05***	-	11.44***
	2016	2		9.70***	-	2.01ns
	2017	1, 2		36.16***	0.39ns	0.25ns
	2018	1		8.87***	-	0.66ns

	2018	2		10.61***	-	0.00ns
DAPR	2016	1		19.53***	-	11.16***
	2016	2		16.77***	-	0.01ns
	2017	1, 2		46.17***	0.23ns	1.25ns
	2018	1, 2		14.21***	0.00ns	0.57ns
FW	2016	1	Full	23.46***	-	-
	2016	1	Reduced	25.21***	-	-
	2016	2	Full	22.17***	-	-
	2016	2	Reduced	32.59***	-	-
	2017	1	Full	19.82***	-	-
	2017	2	Full	9.05***	-	-
	2017	1, 2	Reduced	6.67***	1.87ns	-
Brix	2016	1	Full	4.55***	-	-
	2016	2	Full	5.40***	-	-
	2016	2	Reduced	6.80***	-	-
	2017	1		7.10***	-	17.12***
	2017	2	Full	3.57***	-	-
	2017	2	Reduced	5.75***	-	-
	2016	1	Reduced	7.91***	-	-
pH	2016	1	Full	8.09***	-	-
	2016	1	Reduced	13.61***	-	-
	2016	2	Full	11.97***	-	-
	2016	2	Reduced	14.15***	-	-
	2017	1	Full	2.80***	-	-
	2017	1	Reduced	8.61***	-	-
	2017	2		16.43***	-	27.48***
CRT	2016	1	Full	3.79***	-	-
	2016	1	Reduced	3.55***	-	-
	2016	2		4.02***	-	1.42ns
	2017	1	Full	2.90***	-	-
	2017	1	Reduced	5.73***	-	-
	2017	2		3.45***	-	0.08ns

* $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, ns = not significant

Chapter 2:

Combining Ability for Yield, Fruit Quality, and Water-Stress Tolerance Traits in Introgression Lines Derived from Wild Tomato (*Solanum habrochaites*)

Abstract

Global climate change is reducing the amount of precipitation that provides fresh water for crop production in arid regions. Cultivated tomato, *Solanum lycopersicum*, is susceptible to water-stress, which reduces yields. A tomato wild relative, *Solanum habrochaites*, is water-stress tolerant and may serve as a source of abiotic stress tolerance traits for breeding tomato. A set of five introgression lines (ILs) derived from *S. habrochaites* were selected and evaluated for their potential use in a hybrid tomato breeding program. The five ILs served as parents in crosses with 3 processing tomato inbred lines in a Design II mating design to obtain 15 F1 hybrids. These genotypes (8 parent lines and their 15 F1 hybrids) were evaluated for 11 traits including yield, fruit quality, and water-stress tolerance-related traits in replicated field experiments using a split-plot experiment design over 3 locations. Water treatments (reduced and full irrigation) were assigned to main plots, and genotypes to subplots. Statistical analyses, including for combining ability, were performed on a per-trait basis. GCA estimates ranged between -52.55 and 75.21 were obtained for each of the five IL parents and three inbred line parents for each of the traits. SCA estimates ranged between -36.59 and 26.86. Heritability estimates of 0 to 0.68 were obtained for the traits analyzed for combining ability. Heterosis coefficients using potency ratio had values between -65.0 and 95.0 for each hybrid for each trait. Best parent heterosis was identified in at least one F1 hybrid for each trait (except for traits DAPG1 and FW2). IL parent LA3933 possessing an introgression on chromosome 4 was

identified as the best candidate for use in a hybrid breeding program. The reduced water treatment saved between 27-58% of total applied irrigation water yet did not decrease yield even in material not known to be water-stress tolerant. This reduced irrigation scheme has the potential to be adopted by processing tomato growers to save irrigation water.

Introduction

Tomato, *Solanum lycopersicum,* is one of the most widely grown vegetable crops globally, second only to potato (*Solanum tuberosum*) (FAOSTAT, 2021). Tomato is also one of the most economically significant vegetable crops in the United States (USDA NASS, 2022). California is responsible for the majority of tomato production in the US, including both fresh market and processing market classes (USDA NASS, 2022). In 2021 California tomato production of all market classes had a farmgate value of 1.18 billion US dollars (USDA NASS, 2022). A total of 10.78 million tons of processing tomatoes were produced in the state in 2021, worth 905 million US dollars.

Climate change is expected to decrease yields of tomatoes produced in California due to rising temperatures and reduced irrigation water availability over the next century (Lee et al., 2010; Diffenbaugh et al., 2015; Pathak et al., 2018; Ray et al., 2020). California has a Mediterranean climate with dry warm summers and cool wet winters (Deitch et al., 2017; Pathak et al., 2018; Seager et al., 2019). During the dry summer months, winter snowpack in the Sierra Nevada mountains is a key source of surface water for irrigation. Climate change has led to a decrease in snowpack in the western United States over recent decades and is predicted to continue to decline with further increases in global temperatures (Mote et al.,

2005; Mahoney et al., 2021). The decrease in precipitation is a threat to freshwater resources for irrigation that is required for crop production, including tomatoes (Hartz et al., 2008; Pathak et al., 2018; Ray et al., 2020).

Cultivated tomato has a limited genetic base for breeding improvement as a result of several genetic bottleneck events that occurred during domestication (Rick 1983; Miller and Tanksley, 1990; Corrado et al., 2013; Kulus, 2018; Tamburino et al., 2020). In contrast, wild tomato species are a rich source of genetic diversity for expanding the limited genetic base of cultivated tomato. Tomato wild relatives have been used as sources of important traits for breeding, including biotic and abiotic stress tolerances (Rick, 1983; Bai, et al., 2018; Schouten et al., 2019). *Solanum habrochaites* is reported to be water-stress tolerant (Rick, 1983; Spooner et al., 2005). *S. habrochaites,* other tomato wild species relatives, and cultivated tomato share a high degree of synteny and are inter-crossable with each other (Chetelat and Ji, 2007; Moyle 2008). The St. Clair lab previously mapped QTL for water-stress tolerance-related traits to chromosome 9 of *S. habrochaites* (Truco et al., 2000; Goodstal et al., 2005; Arms et al., 2015; Lounsbery et al., 2016; Groh et al., 2022). In addition to *S. habrochaites*, other tomato wild relatives including *S. lycopersicum* var. *cerasiforme*, *S. pennelii*, *S. chilense*, and *S. sitiens* were reported as tolerant to water-stress (Rick 1973; Chetelat et al., 2009; Moyle and Muir, 2010; Dariva et al., 2020).

Water-stress tolerance-related traits can be evaluated on plants grown under deficit irrigation (Payero et al., 2009; Richards 2006; Tuberosa 2012). In contrast to drought, deficit irrigation is the continued application of water throughout a growing season that is less than the total evapotranspiration needs of a crop. Direct measurement of water-stress tolerance of a

crop in a field situation is difficult, laborious and technologically challenging. As an alternative, proxy traits such as yield, harvest index, shoot biomass, and carbon isotope discrimination are commonly used to indirectly measure water-stress tolerance (Payero et al., 2009; Richards 2006; Tuberosa 2012).

Agriculturally important traits for processing tomatoes include yield, maturity, plant size, fruit soluble solids (Brix), pH, fruit size, and fruit weight (Barrett et al., 2007; Hartz et al., 2008; Barrios-Masias and Jackson, 2014). Tomato growers are paid by food processors primarily based on tonnage (yield) (Hartz et al., 2008). To facilitate end of season mechanical harvest, processing tomatoes must have uniform fruit maturity, adequate fruit size and weight, and compact plant size (Hartz et al., 2008; Barrios-Masias and Jackson, 2014). Fruit pH must also be below 4.4 to maintain food safety in canned products (Anthon et al., 2011).

S. habrochaites possesses beneficial horticultural and fruit quality traits in addition to abiotic and biotic stress tolerances, including tolerance to root chilling, resistance to tomato yellow leaf curl virus, and resistance to herbivores including insects (Glas et al., 2012; Rick 1973; Truco et al., 2000; Yang et al., 2014). Introgression lines (ILs) derived from *S. habrochaites* exhibited increased fruit soluble solids content (measured in degrees Brix) compared to the recurrent parent (Bernacchi et al., 1998; Kubond and St. Clair, 2022). An IL library is a set of genotypes that each contain a different unique introgression from a wild species donor parent in the background of a cultivated recurrent inbred parent. IL libraries for various tomato wild relatives have been created (Eshed and Zamir, 1994(a); Fulton et al., 2000; Canady et al., 2005; Chetelat et al., 2019). Introgression lines from *S. habrochaites* contain alleles that contribute to

increased average fruit weight, reduced fruit sunburn, and increased plant canopy cover (Kubond and St. Clair, 2022).

Similar to some other commercially bred vegetable crops, F1 hybrids are the most common type of cultivar used in processing tomato production in the US; nearly all commercial processing tomato varieties are hybrids (Janick, 1998; Wehner 1999; Hartz et al., 2008; Processing Tomato Advisory Board, 2021). F1 hybrid cultivars are created by first developing inbred parent lines, then crossing pairs of lines with each other through controlled pollinations to obtain F1 hybrids (Fehr 1987; Rehman et al., 2021). Subsequently, the hybrids are evaluated for traits of importance and compared to each other to determine the best combination of inbred parents that produce agriculturally superior performing F1 hybrids.

Evaluation of combining ability on a per trait basis can be used to identify and select inbred parents for producing the best performing F1 hybrid combinations. A mating design, such as Design II, creates a structured population that facilitates estimation of general combining ability (GCA), specific combining ability (SCA), and the relative magnitude of additive and nonadditive genetic variance on a per trait basis (Comstock and Robinson, 1948). Design II has been employed in tomato studies to determine GCA and SCA of various traits (Hartman and St. Clair, 1999; Haggard and St. Clair, 2014; Liu et al., 2021). It has also been applied in breeding various crops including corn, cotton, oat, and sorghum(Wang et al., 1999; Zeng et al, 2011; Holland and Munkvold, 2001; Godoy and Tesso, 2013; Hirut et al., 2017). Data obtained in combining ability experiments can be used to obtain estimates for heterosis and heritability (Fehr 1987; de Vienne and Fievet, 2020)

The objective of this study was to determine which IL parents could be suitable for potential use in a F1 hybrid cultivar breeding program for the improvement of horticultural, fruit quality, and water-stress tolerance-related traits in processing tomato.

Materials and Methods

Plant Material

A total of eight parental inbred (homozygous) cultivated tomato lines were used in this study. Five parents were designated as females and three parents were designated as males. The five female parents (LA3913, LA3933, LA3938, LA3956, and LA3968) were tomato introgression lines (ILs) that each contain a single unique chromosome introgression from wild *S. habrochaites* accession LA1777 in the genetic background of *S. lycopersicum* processing tomato inbred cultivar E6203 (Monforte and Tanksley, 2000(a)). Seeds of the ILs were obtained from the C.M. Rick Tomato Genetics Resource Center at UC Davis (tgrc.ucdavis.edu). The three male parents (Apex 1000, Heinz 8761, and CTRI 1558) are processing tomato inbred lines were selected to represent some of the diversity in California processing tomato germplasm (Park et al., 2004). Seeds of the three inbred lines were obtained from stocks maintained by the St. Clair tomato breeding program at UC Davis.

The five female parents were selected from a larger set of 24 *S. habrochaites*-derived ILs that were evaluated in 2016 and 2017 in replicated field trials (data not shown; Kubond and St.Clair, 2022) using a base selection index (Williams, 1962). Selection using a base index involves assigning an economic weight to each trait and the phenotypic value for each of the traits (Williams, 1962; Geidel et al., 2000; Marulanda et al., 2021). The selection index used to select IL parents included three traits, Brix, ripe fruit yield (RYLD), and harvest index (HI), which

were each assigned equal economic weights (see Phenotyping Traits, below). ILs LA3913, LA3933, LA3938, LA3956, and LA3968, each containing a single introgression on chromosomes 1, 4, 5, 9, and 12, respectively, were chosen for this study.

Plants of the eight parental lines (i.e., five ILs and three inbred lines) were grown in 3L pots in a UC Davis greenhouse using standard tomato cultivation practices. Self-pollinated seed from each line was obtained during Fall 2017 in sufficient quantities for replicated field experiments. A Design II mating design was also used in a controlled cross-pollination scheme among potted parental plants to obtain seeds of 15 hybrids for use in replicated field trials. Cross pollinations were performed during Fall 2017-Winter 2018 and Fall 2018-Winter 2019. Sufficient hybrid seed was obtained for all 15 hybrid combinations for summer 2018 experiments. For summer 2019 experiments, sufficient hybrid seed was obtained for 12 of 15 hybrids. Three hybrids yielded insufficient seed quantities so were not included in 2019: LA3938 x CTRI 1558, LA3913 x Heinz 8761, and LA3913 x CTRI 1558.

Field Experimental Design

Field experiments were conducted during the summers of 2018 and 2019 at the University of California Davis Plant Sciences Field Research Facility in Davis, California. California has a Mediterranean climate that is characterized by cool wet winters and dry warm summers (Deitch et al., 2017; Pathak et al., 2018; Seager et al., 2019). During the summer field experiments in Davis, no measurable rain fell, and the plants only received water via irrigation.

Two field locations were used in 2018 and one in 2019. Location for this experiment was defined as fields with different cropping histories. A split plot experimental design was

employed. Each location contained two full repetitions of the split plot. Each split plot had two main plots, assigned to two different irrigation treatments (full crop evapotranspiration and reduced, see details below). Main plots consisted of 4 blocks of subplots, with subplots assigned to genotypes (15 hybrids, 8 parent lines, and control E6203). Each subplot consisted of 8 plants. A single row of plants was planted on each 154.2 cm center bed, with 30.5 cm within-row spacing between plants. A within-row alley of 91.4 cm without plants was included between the end and beginning of each subplot within each row to facilitate access for data collection. Double border rows of various processing tomato inbred line cultivars were planted between each main plot and on the outside perimeter of the field experiment to minimize edge effects.

Seeds of each genotype were seeded into flats in a UC Davis greenhouse during spring of 2018 and 2019 and grown using standard horticultural practices for tomato. Once plants reached the second true leaf stage (approximately 5 weeks), flats were transferred to a lath house to harden off for a week prior to transplanting to the field plots by hand.

Two drip irrigation treatments were used: normal water or full crop evapotranspiration (ETc) of tomato, and a reduction post-full fruit set to 40% ETc for the remainder of the growing season. The reduced water treatment in this study was selected due to the possibility of adoption by California processing tomato growers to save water if it did not lead to significant yield reduction. Full fruit set was defined as 51% of E6203 control plots scored with a first ripe fruit. The reduced water treatment was applied after full fruit set was achieved to limit impacts on floral abortion and subsequent fruit yield. Tomato can experience floral abortion due to water-stress (Ruan et al., 2012). The amount of water applied to the reduced water treatment

was determined by multiplying the amount of water to apply to the full water treatment by 0.4. All subsurface drip irrigation water was precisely applied and the amount was calculated based on canopy width measurements and ETc.

Overhead sprinkler irrigation was applied for two weeks post-transplant to enable root establishment prior to switching to drip irrigation. Subsequently, water was applied to plants via 1.58 cm width drip tape per bed (row), with 30.48 cm emitter spacing (Toro Flow Control) sub-surface drip irrigation, buried 20 - 22 cm below the soil surface of each bed prior to transplant. Irrigation water was applied three times weekly to evenly space the water application over weekly periods. The amount of water to apply was determined by the amount of evapotranspiration (ET), as recorded by a nearby CIMIS station (cimis.water.ca.gov) since the prior irrigation, and the canopy size of the tomato plants. The water flow rate was calculated prior to each irrigation due to daily fluctuations in water pressure. The length of each irrigation treatment was calculated using the water flow rate of the irrigation system and the calculated amount of water to apply to each treatment as determined by ETc for tomato. The irrigation system was manually turned on and off for each irrigation period. The amount of water applied to each treatment and location combination was recorded using water flow meters (Sensus VMSR2 Brass Water Meter). During the field experiments, no measurable summer precipitation was recorded. Urea nitrogen fertilizer was banded prior to transplant at a depth of 15-20 cm beneath the surface of each bed at a rate of 22.41 Kg/ha (Hartz et al., 2008). Urea nitrogen fertilizer was applied via fertigation for four weeks beginning 3 weeks after transplanting, at a rate of 33.6 Kg/ha, starting 21 days post-transplant. A total of 156.8 Kg/ha of Urea nitrogen fertilizer was applied through the duration of the season.

Phenotyping Traits

For all three locations, trait data was collected on the subplots on a per-plot basis. The traits were grouped into the following categories: horticultural, fruit quality, and water-stress tolerance-related (Table 1).

Table 1: Traits evaluated by category on a per-plot basis.

Trait Category	Trait Code	Description
Water-Stress Tolerance-Related	SDW	Shoot dry weight (biomass, in g)
	SFW	Shoot fresh weight (biomass, in g)
	BioM	Total fresh above ground biomass at harvest (g)
Horticultural	RYLD	Ripe fruit yield (kg)
	TYLD	Total fruit yield (kg)
	DAPG	Days after planting to first green fruit
	DAPR	Days after planting to first ripe fruit
	HI	Harvest index (percentage of ripe fruit to total biomass)
Fruit Quality	FW	Weight (g) of 25 fruits
	pH	pH of tomato fruit slurry
	Brix	Soluble solids content of tomato fruit slurry (degrees Brix)

Days after planting to first green fruit (DAPG) was defined as the number of days from transplanting until greater than 50% of the plants in a subplot each had a green fruit of at least 1 cm in diameter (Lounsbery et al., 2016). Scoring was conducted three times a week, on Mondays, Wednesdays, and Fridays. Days to first red fruit (DAPR) was defined as the number of

days from transplanting until greater than 50% of the plants in a subplot had at least one fully ripe fruit (Lounsbery et al., 2016). As with DAPG, DAPR was scored three times weekly on Mondays, Wednesdays, and Fridays.

Subplots were destructively harvested once all the plants in a subplot had approximately 90-95% ripe fruit, or if a subplot was late maturing relative to E6203, at the end of the growing season regardless of percentage of ripe fruit load. Two plants from each subplot were cut at the soil line. Fruit was removed manually from the two shoots per subplot, separated into fully ripe and not ripe, and each group was weighed to obtain total fruit yield and ripe fruit yield (Lounsbery et al., 2016). Subsequently, the two shoots were placed into an onion mesh bag and weighed immediately to obtain fresh shoot weight. Harvest index was calculated using the following formula: HI = RYLD/(TYLD + SFW). Total fresh biomass was determined by adding total fruit yield to shoot fresh weight. Harvested tomato shoots in the mesh bags were placed in a forced air dryer and allowed to dry for at least two weeks, then weighed to obtain shoot dry weight per subplot (Lounsbery et al., 2016).

Once each subplot had sufficient numbers of ripe fruit based on of visual inspection of subplots, 25 ripe fruit were selected randomly and harvested. The 25 fruit were weighed and the weight divided by 25 to obtain the average fruit weight (FW) per subplot. After weighing, each fruit was cut in half longitudinally and one half from each of the 25 fruits were blended together for 1 minute. The resulting tomato slurry was measured immediately for Brix using a digital hand-held pocket refractometer (Atago LTD) and pH using a portable pH meter (Oakton pH 150).

Trait data was collected for each parent line at all three locations, and for all 15 hybrids at locations 1 and 2. At location 3, trait data was obtained for 12 of the 15 hybrids. Due to insufficient seed, three hybrids (LA3913 x CTRI 1558, LA3938 x CTRI 1558, and LA3913 x Heinz 8761) were not included in location 3.

Statistical Analysis

Statistical analysis was performed on a per trait basis. Prior to combining ability analysis, analysis of variance (ANOVA) was performed for each trait using the R stats package (R Core Team) and lme4 (Bates et al., 2015). Each trait dataset was checked for normality with Shapiro-Wilk W-statistic and a Quantile-Quantile plot (R Core Team). Homogeneity of variance was evaluated using residuals plots (R Core Team).

Data for each trait was pooled across each of the three locations initially for analysis. Significant location x genotype interactions were detected for each trait except DAPR. Consequently, ANOVA for each trait (except DAPR) was performed separately for each location using the following linear additive model with the lmer function in the lme4 package in R (Bates et al., 2015):

Trait = Water + Genotype + (Water x Genotype)

For DAPR the following linear additive model was used:

Trait = Loc + Water + Genotype + (Water x Loc) + (Genotype x Loc) + (Genotype x Water) + (Genotype x Water x Loc)

Loc refers to the field location (loc1, loc2, or loc3), Water refers to the water treatment (Full ETc or 40% ETc post-fruit set), Genotype refers to the 8 parent lines, 15 hybrids, and E6203. When significant ($P \leq 0.05$) genotype x environment (G x E) interactions were detected for a

particular trait dataset, analysis was performed by location, treatment, or combination depending on the source of variation revealed in an interaction.

Data transformations were performed for RYLD1F, RYLD1R, RYLD2F, RYLD2R, TYLD1F, TYLD1R, TYLD2, and DAPG1 to meet ANOVA assumptions. A square root transformation was applied to RYLD1F, RYLD1R, RYLD2F, RYLD2R, TYLD1F, TYLD1R, and TYLD2 and an inverse transformation (1/x) was used for DAPG1. Despite using a transformation, DAPG1 still did not meet the assumption of normality, but did meet the assumption of homogeneity of variance. No attempted transformation improved both normality and homogeneity of variance simultaneously for DAPG1, likely due to a large number of plots reaching the same maturity stage on the same day. An inverse transformation was selected for DAPG1 due to its improvement in reducing heterogeneity of variance as determined by inspection of residual plots.

Means separations were performed with Tukey's, which compared each genotype to each of the other genotypes with the cld(contrasts) function in the lsmeans package (Lenth et al., 2016). When significant G x E interactions (genotype x location, or genotype x water treatment) were detected, trait data is referred to by their specific location and treatment combination. For example, total fruit yield (TYLD) from location 1 under the full water treatment is denoted as TYLD1F, with the location indicated after the trait code first, either 1 for location 1 in 2018, 2 for location 2 in 2018, or 3 for the 2019 location, followed by the water treatment, F for full water treatment or R for reduced water treatment.

For each trait that displayed significant differences among genotype means using ANOVA, the IL parents, inbred line parents, and their F1 hybrids were analyzed for combining

ability using a Design II mating design ANOVA (Comstock and Robinson, 1948; Godoy and Tesso, 2013; Haggard et al., 2014; Hirut et al., 2017). Estimates for GCA and SCA were obtained using REML method with the mmer function in R sommer package (Covarrubias-Pazaran 2016; Covarrubias-Pazaran, 2022). For each trait, random effects and fixed effects were specified in the model. The following model was used to obtain estimates for GCA and SCA:

Trait = Loc + Water + (Female*Male) + Female + Male + (Female*Male*Loc) + (Female*Loc) + (Male*Loc) + (Female*Male*Water) + (Female*Water) + (Male*Water) + (Water*Loc)

Loc refers to the field location, Water refers to the water treatment, Female is the maternal effect on progeny performance (GCA), Male is the paternal effect on progeny performance (GCA), and Female*Male is the interaction of maternal and paternal effects (SCA). Terms in parentheses in the model were considered random. The function randef() was used to obtain female (GCA), male (GCA), and female*male (SCA) effects. When significant statistical interactions between main effects were detected in ANOVA, trait data was separated and analyzed separately as appropriate, depending on the type of interaction detected.

Heritability estimates for both narrow (h^2) and broad sense heritability (H) were determined for each trait. Variance components were calculated as part of the combining ability analysis. *F* is Wright's inbreeding coefficient. Additive variance for both males and females were calculated as Va(female) = (4/(1+*F*)) * Vfemale where *F* = 1 due to complete inbreeding (Hallauer and Miranda, 1981; Covarrubias-Pazaran, 2022). Estimates for dominance variance were calculated as Vd = (4/(1+*F*)^2) * Vfemale*male where *F* =1 due to complete inbreeding. To determine the proportion of genetic variance from additive genetic variance, the

ratio of h^2 to H was calculated on a per trait basis. An h^2/H ratio of 1 indicates that all genetic variance is from Va. An h^2/H ratio of 0 indicates that all genetic variance is from Vd and/or Vi.

Heterosis coefficients relative to the mid-parent were calculated for each trait for which significant differences between genotype means were detected. Potence Ratio (H_{PR}) was used to calculate heterosis coefficients (de Vienne and Fievet, 2020). H_{PR} was selected over other heterosis coefficients due to H_{PR} including both parental trait means as well as the trait means for the F1 hybrids. H_{PR} was calculated as:

$$H_{PR} = Z_{12} - \bar{Z}/((Z_2-Z_1)/2)$$

Where Z_{12} is the phenotypic value of the hybrid, Z_1 is the phenotypic value of the first parent, Z_2 is the phenotypic value of the second parent, and $\bar{Z}$ is the mean parental phenotypic value. Additivity indicates that for a particular trait, $Z_{12} = \bar{Z}$ (de Vienne and Fievet, 2020).

Results

Significant genotype effects ($P \leq 0.05$) in ANOVA were detected for most traits (except for SFW1, SFW2, SDW2, and BioM2) (Supplemental Table 1). Significant genotype effects were detected for RYLD3, TYLD2, SFW3, SDW2, and BioM1 using ANOVA, but trait means for each genotype were not significantly different from each other with Tukey's (Supplemental Table 2). Traits lacking either a significant genotype effect or significant difference among traits means for each genotype were not analyzed for combining ability, heritability, or heterosis (Kearsey and Pooni, 1996; Haggard et al., 2014).

Significant ($P \leq 0.05$) genotype x location interactions were found for each trait except DAPR (Supplemental Table 1). Subsequent analysis for each trait (except DAPR) was performed separately for each location. In addition, significant genotype x treatment interactions were

detected for RYLD1, RYLD2, TYLD1, FW3, and pH1, and subsequent analysis for each of these traits was performed separately for each combination of location and treatment.

General Combining Ability:

GCA estimates were obtained for each of the five IL parents and the three inbred line parents for each trait that combining ability analysis was performed (Supplemental Table 3). Among the IL parents there were mixed results for GCA values. None of the IL parents had GCA estimates that were horticulturally positive for most or all traits. LA3933 had the largest number of traits with horticulturally positive GCA estimates while LA3938 had the fewest. The remaining three IL parents were intermediate between LA3933 and LA3938 for the number of traits with horticulturally positive GCA estimates. Nearly all of the water-stress tolerance-related traits lacked either significant genotype effects or significant differences among the trait genotype means (Supplemental Table 1; Supplemental Table 2).

LA3933 with a chromosome 4 introgression from *S. habrochaites* had the largest number of traits with GCA estimates that were horticulturally positive among the IL parents (Supplemental Table 3). Notably, LA3933 had GCA estimates with positive horticultural value for RYLD1F, RYLD1R, and TYLD1F and GCA estimates of zero for the remaining yield traits (Table 2). It was the only IL parent that did not exhibit horticulturally undesirable estimates for GCA for any yield related traits. Despite LA3933 having no horticulturally undesirable estimates for GCA for any of the yield related traits, GCA estimates of zero obtained for RYLD2F, RYLD2R, and TYLD3 indicate that for these yield traits, LA3933 may not be a favorable parent.

Table 2: GCA estimates for yield traits for each female parent (ILs) and male parent (tomato inbred cultivars). For each trait, trait means and Tukey groups are listed. If trait data required transformation, the untransformed trait mean is included in parenthesis next to the transformed trait mean. GCA estimates for each trait are listed beside the trait mean and Tukey group. A higher GCA estimate indicates better combining ability horticulturally. RYLD2F, RYLD2R, TYLD3 had GCA estimates of zero for each parent (Supplemental Table 3) so are not included in this table.

Females	RYLD1F	RYLD1F GCA	RYLD1R	RYLD1R GCA	TYLD1F	TYLD1F GCA	TYLD1R	TYLD1R GCA
LA3913	2.41 (5.80)a-b	0.027	2.26 (5.12)a	-0.105	2.76 (7.64)a-b	0.083	2.56 (6.53)a-b	-0.076
LA3933	2.72 (7.39)b	0.019	2.36 (5.57)a-b	0.014	2.98 (8.86)a-b	0.013	2.43 (5.92)a-b	0.000
LA3938	2.66 (7.07)a-b	0.005	2.35 (5.5)a-b	-0.203	2.91 (8.47)a-b	0.005	2.46 (6.04)a-b	-0.150
LA3956	2.90 (8.39)b	-0.027	2.39 (5.73)a-b	0.117	3.26 (10.66)b	-0.066	2.54 (6.46)a-b	0.098
LA3968	2.79 (7.79)b	-0.025	2.75 (7.56)a-b	0.177	2.92 (8.52)a-b	-0.034	2.85 (8.1)a-b	0.127
Males								
Apex 1000	2.55 (6.51)a-b	0.113	2.56 (6.55)a-b	0.042	2.82 (7.94)a-b	0.116	2.73 (7.47)a-b	9.811E-03
CTRI 1558	1.89 (3.58)a	0.009	2.05 (4.19)a	0.011	2.18 (4.76)a	-0.055	2.20 (4.82)a-b	7.769E-05
Heinz 8761	2.52 (6.37)a-b	-0.122	2.21 (4.88)a	-0.053	3.06 (9.38)a-b	-0.061	2.65 (7.03)a-b	-9.890E-03

Specific Combining Ability:

Estimates for SCA were determined for each F1 hybrid progeny for each trait except for those traits that either had a non-significant genotype effect or no significant differences among genotype means (Supplemental Table 4). LA3956 x Apex 1000 had the largest number of traits (15 of 24) with horticulturally positive SCA estimates. This same hybrid exhibited horticulturally positive SCA estimates for each of the yield traits except TYLD3. The other hybrids showed horticulturally undesirable SCA estimates for two or more yield-related traits.

There was not a discernable pattern of any IL parent contributing to hybrids with horticulturally positive SCA estimates when combined with the inbred line parents. LA3913 combined with Apex 1000 or CTRI 1558 had hybrids with a larger number of traits with horticulturally positive SCA estimates, but not when combined with Heinz 8761. This observation was similar for the other IL parents for which a higher number of horticulturally positive SCA estimates were detected when crossed with one or two of the inbred line parents, but not all three.

Heritability Estimates:

Heritability estimates for both narrow sense (h^2) and broad sense heritability (H) were obtained for each trait for which significant genotype effects were detected (Supplemental Table 5). Both H and h^2 values ranged between zero and 0.682. Values obtained for the h^2/H ratios for each trait (Supplemental Table 5) ranged from zero to 1. An h^2/H ratio value of zero indicates that the genetic variance for a particular trait is entirely 1 from dominance and/or

epistatic variance. Ratio values of 1 indicate that all of the genetic variance for a particular trait is due to additive genetic variance (Hallauer and Miranda, 1981).

In general, traits in the fruit quality category and HI tended to have higher h^2 and H estimates. Yield and water-stress tolerance-related traits tended to have lower estimates for h^2 and H. Additionally, fruit quality traits tended to have a higher proportion of additive variance than the other categories of traits, with h^2/H ratios from 0.79 to 1. Water-stress tolerance-related traits had the lowest h^2/H ratios of zero for both SDW3 and BioM3. The H estimates for these traits were also very low at 0.099 and 0.07, respectively. Yield-related traits showed the largest range of h^2/H ratio values, between 0 and 1.

Heterosis:

Heterosis coefficient H_{PR} was calculated for each hybrid for each trait that had significant genotype effects and significantly different means among genotypes (Figure 1). Best parent heterosis was found for at least one of the hybrids for each trait, except DAPG1 and FW2. Each hybrid exhibited best parent heterosis for more traits than worst parent heterosis (WPH). Among the hybrids, LA3968 (chromosome 12 introgression) x CTRI 1558 and LA3913 (chromosome 1 introgression) x Apex 1000 had the largest number of traits (17) with best parent heterosis.

	RYLD1F	RYLD1R	RYLD2F	RYLD2R	TYLD1F	TYLD1R	TYLD3	HI1	HI2	HI3	DAPG1	DAPG2	DAPG3	DAPR	SDW3	BioM3	FW1	FW2	FW3F	FW3R	Brix1	Brix2	Brix3	pH1	pH2F	pH2R	pH3
Hybrid:																											
LA3938 x Heinz 8761	-1.0	2.9	5.3	1.7	-4.2	1.0	0.1	1.3	-2.1	-3.7	-0.1	-0.2	-1.1	-7.4	-0.9	0.2	-1.5	-1.5	-0.1	2.0	0.5	-0.1	-1.3	-1.4	-1.8	-5.0	NaN
LA3938 x CTRI 1558	1.3	-1.3	3.2	1.8	0.9	-2.2		2.0	2.2		17.0	-6.6					0.2	-0.2			5.5	2.1		-2.0	-5.0	-3.0	
LA3938 x Apex 1000	8.3	2.2	10.1	3.0	7.9	1.6	0.7	5.9	7.0	0.5	-1.0	-8.0	-3.7	-1.6	-1.3	0.7	0.9	-6.6	0.8	0.8	0.0	1.3	0.8	-2.3	-3.0	-1.0	NaN
LA3913 x Heinz 8761	6.8	-0.2	95.0	1.6	1.6	-6.3		4.6	1.4		0.3	-0.5					-2.9	-1.2			0.3	1.5		-0.9	-0.9	-1.0	
LA3913 x CTRI 1558	2.8	6.6	8.9	9.8	1.8	3.4		5.2	3.7		4.0	-10.4					0.9	0.4			1.8	3.5		-0.9	-0.4	-1.6	
LA3913 x Apex 1000	6.7	1.1	5.2	0.4	16.7	0.3	1.5	3.4	2.8	-1.5	5.0	-65.0	-0.2	-1.3	0.6	1.7	8.2	0.5	4.4	2.6	1.2	3.6	0.0	-0.9	-1.1	-1.4	-0.7
LA3956 x Heinz 8761	-1.6	5.6	0.2	-0.2	-5.5	7.0	-0.4	0.1	0.1	-1.6	0.6	-2.7	-0.7	-0.1	0.3	-0.3	-1.3	-3.8	-6.3	1.5	0.6	3.3	10.0	-1.0	1.0	0.5	-1.0
LA3956 x CTRI 1558	0.6	3.6	2.6	0.2	0.0	0.4	3.4	1.1	1.0	0.7	6.6	-3.4	-1.2	-1.3	0.1	5.9	0.5	-1.9	0.5	0.5	0.8	1.2	1.3	-0.5	0.3	-0.4	-0.6
LA3956 x Apex 1000	1.2	6.5	6.0	-0.6	0.0	5.5	-1.2	1.0	1.0	-17.3	5.0	-4.1	-2.2	-1.0	-1.0	-1.0	0.1	-0.9	4.2	2.1	4.1	6.0	3.3	-1.0	-0.3	0.0	-1.5
LA3968 x Heinz 8761	-1.5	1.1	4.6	-0.3	-1.4	0.9	-0.5	0.9	0.6	0.8	0.7	0.0	-1.4	-4.3	-0.7	-0.5	-0.1	-1.2	0.3	1.4	0.3	1.4	-1.5	NaN	NaN	-15.0	-5.0
LA3968 x CTRI 1558	1.2	2.4	1.6	0.4	1.2	2.5	5.2	1.0	0.9	2.2	3.7	-5.1	-0.2	-2.0	-0.1	20.2	-0.3	-0.8	1.2	0.6	4.4	2.9	-0.8	-1.7	-4.3	-1.6	-1.9
LA3968 x Apex 1000	-0.9	3.0	15.7	1.8	-4.4	3.7	-0.3	0.7	0.9	-11.0	3.0	-0.6	-5.8	-0.8	-0.1	-0.2	-0.5	0.0	0.3	0.7	0.0	0.2	-0.5	-2.0	-1.7	-1.4	-5.0
LA3933 x Heinz 8761	1.8	2.2	7.9	3.2	-2.0	-0.1	-0.8	1.2	0.2	-1.5	0.1	-1.0	1.0	-15.7	-1.2	-0.8	0.1	-0.1	0.2	0.6	0.1	0.1	-1.1	-1.0	-1.7	-3.0	0.1
LA3933 x CTRI 1558	1.1	3.6	4.6	2.1	0.6	4.1	10.7	1.8	1.5	1.4	3.0	-1.2	0.8	-3.0	0.9	6.7	0.5	0.2	0.9	0.8	-1.0	1.1	-1.6	-0.1	-1.7	-2.0	-1.0
LA3933 x Apex 1000	4.3	4.8	0.0	23.0	3.5	3.1	1.0	1.2	1.9	-3.6	1.0	-2.0	0.9	-1.8	0.3	1.0	0.5	0.4	0.4	0.4	1.6	0.9	-0.8	-0.3	-0.7	-1.0	-0.4

Figure 1: Heterosis (potency ratio) coefficients for 15 tomato F1 hybrids evaluated under two water treatments. Hybrids are listed to the left by their parental combinations. H_{PR} coefficients are listed for each trait. Green shaded cells indicate best parent heterosis, blue cells indicate best parent value, gray cells indicate additivity, yellow cells indicate negative mid parent value, and red cells indicate worst parent heterosis. Heterosis coefficients were interpreted in terms of horticultural desirability. Blank cells indicate the hybrid was not included in that location for that trait. Cells with NaN indicate that H_{PR} coefficient for that hybrid was a non-existent number since the denominator contained a zero ($H_{PR} = Z_{12} - \bar{Z}/((Z_2 - Z_1)/2)$). Blank cells indicate that the hybrid was not present at a particular location.

Among yield traits, each of the hybrids (except LA3956 x Heinz 8761 and LA3968 x Heinz 8761) exhibited best parent heterosis for more traits than worst parent heterosis (Figure 1). LA3956 x Heinz 8761 and LA3968 x Heinz 8761 had an equal number of traits with best parent heterosis as WPH, with two each.

SDW3 and BioM3 were the only two water-stress tolerance-related traits that were analyzed for heterosis (Figure 1). Best parent heterosis was detected in two hybrids only, LA3933 x Heinz 8761 and LA3938 x Apex 1000 for SDW3 with H_{PR} values of 1.2 and 1.3 (Figure 1). For BioM3, 4 hybrids showed best parent heterosis: LA3913 x Apex 1000, LA3933 x CTRI 1558, LA3956 x CTRI1558, and LA3968 x CTRI1558 with H_{PR} values of 1.7, 6.7, 5.9, and 20.2 respectively (Figure 1).

Discussion

General Combining Ability for Yield

LA3933 is the only IL parent in this study that may be suitable for use in a processing tomato hybrid breeding program to improve certain key traits including yield, harvest index, and maturity (Barrett et al., 2007; Hartz et al., 2008; Barrios-Masias and Jackson, 2014) because it has predominantly horticulturally positive GCA estimates for TYLD, RYLD, HI, and DAPG (Supplemental Table 3). Each of the other four IL parents exhibit GCA estimates in at least one location that suggest those parents would have detrimental horticultural effects on TYLD, RYLD, HI, and DAPG. Employing LA3933 to improve TYLD, RYLD, HI, and DAPG should be weighed against its negative horticultural effects as a parent on FW and Brix (Supplemental Table 3). Evaluation of LA3933 for a particular breeding program could be assessed with a selection index to weigh traits by their relative economic importance (William, 1962; Jahufer and Casler, 2015; Beyene et al., 2016). A selection index can help determine if the positive effects that LA3933 has on yield, harvest index, and maturity outweigh its negative on FW and Brix for potential use as a parent in a hybrid cultivar breeding program. Previously, a QTL for Brix was identified in the

same region from *S. habrochaites* that is possessed by LA3933 (Bernacchi et al.,1998). Additionally, further trialing of LA3933 and its hybrid progeny at multiple locations would be required to determine the stability of GCA estimates for LA3933 across environments (Longin et al., 2012; Tadesse et al., 2013; Ginkel and Ortiz, 2018).

Previous combining ability studies in tomato have reported GCA estimates for total yield that were entirely or predominantly non-zero estimates (Bhatt et al., 2001; Liu et al., 2021). GCA is primarily associated with additive genetic variance (Sprague and Tatum 1942; Hallauer and Miranda 1981; Badu-Apraku et al., 2013; Beyene et al., 2017). In our study, GCA estimates of zero were identified for RYLD2F, RYLD2R, and TYLD3, which suggests involvement of dominance, epistatic variance, a combination of negative alleles, or a combination of each. Additionally, as found in Kubond and St. Clair (2022), the large introgressions contained on each IL harbor a number of horticulturally undesirable traits. The differences between GCA estimates for yield among these studies is unsurprising because estimates are dependent on and applicable to the specific inbred parents and their hybrid progeny (Sprague and Tatum 1942; Hallauer and Miranda 1981; Longin et al., 2012).

Combining ability studies in other crops such as corn (Rice and Tracy 2013), cotton (Zhang et al., 2016), and rice (Thalapati et al., 2015) have included wild species or wild species introgression lines in their evaluations of yield. Multi-location trialing was used to evaluate GCA for yield in corn, estimates which varied by location (Makumbi et al., 2011; Rice and Tracy 2013; Qi et al., 2013). In our study, GCA estimates for IL parents were also not consistent across the three locations tested, and H ranged from zero to 0.474, indicating no to moderate heritability

(Supplemental Table 5). Our study provides evidence of the influence of environment on yield as the majority of phenotypic variance is from the environment or G x E effects.

Stability of Performance for Fruit Quality Traits

GCA estimates for each of the ILs were more consistent across locations for one or more of the fruit quality traits (Supplemental Table 3). The relative consistency of GCA estimates suggests that G x E interactions have less effect on fruit quality traits in this set of lines and at these particular locations. Further field trialing would be required to determine if estimates remained relatively stable across additional locations (Eshed et al., 1996; Tanksley et al., 1996; Makanda et al., 2010; Das et al., 2019). However, the use of these ILs to improve these specific traits would be unadvisable. The ILs in this study either do not exhibit consistent GCA estimates across locations for improved horticultural value for fruit quality traits or show horticulturally undesirable GCA estimates for traits such as yield or maturity. Fruit quality traits in tomato would not be weighed more heavily than yield in a selection index since yield is the most economically important trait (Hartz et al., 2008; Barrios-Masias and Jackson, 2014).

Previous findings in tomato by Liu et al. (2021) determined that there was a range of GCA values among parental lines for fruit quality traits including FW. As mentioned previously, a selection index could be employed to weight traits according to their economic value to help select parents for potential use in a hybrid cultivar breeding program (William, 1962; Jahufer and Casler, 2015; Beyene et al., 2016). However, a selection index alone is not sufficient to select superior parents of hybrids (Fehr 1987; Longin et al., 2012). Hybrid progeny of potential

parents must also be evaluated to determine which combination of parents yields the best hybrids in testcrosses for a hybrid breeding program.

In prior studies of combining ability for Brix in tomato, interactions between genotype and location were not detected or only a single location was evaluated (Bhatt et al., 2001; Figueiredo et al., 2015; Figueiredo et al., 2016; Liu et al., 2021). Similar to what was found in this study, not all parents in each of these prior studies had horticulturally desirable estimates for GCA for Brix. GCA estimates for Brix ranged from horticulturally desirable to horticulturally undesirable. In these four prior studies, none of the parental lines were introgression lines and were not known to contain wild species introgressions outside of known disease resistance loci.

Compared to corn and other major agronomic crops, there are few published studies on combining ability in tomato. Commercial (private sector) tomato breeding programs that breed for hybrid cultivars rarely publish any details about their breeding programs (Wehner 1999; Haggard et al., 2014). In addition, limited public research has been published on the use of mating designs to select superior inbred parents for tomato hybrid cultivar breeding (Hartman and St. Clair, 1999; Bhatt et al., 2001; Haggard et al., 2014; Figueiredo et al., 2015; Figueiredo et al., 2016; Liu et al., 2021).

Identification of Best Parent Heterosis in F1 Progeny

Best parent heterosis for a trait is when the F1 progeny has a more horticulturally desirable phenotype than what would be predicted based on the phenotype of the parents (de Vienne and Fievet, 2020). Best parent heterosis was identified in our study for at least one hybrid for each trait, except DAPG1 and FW2 (Figure 1). Previous studies in tomato reported

that heterosis may be present for certain traits, including yield and maturity (Griffing, 1990; Wehner, 1999; Monforte and Tanksley 2000(b); Krieger et al., 2010; Liu et al., 2021). In contrast to genetically narrow cultivated tomato, wild tomato species have a wealth of genetic diversity for use in crop improvement (Rick 1973; Bauchet and Causse, 2012; Rothan et al., 2019). The heterosis detected in this study may be a result of using ILs derived from wild *S. habrochaites*, which are genetically distant from cultivated tomato (Spooner et al., 2005; Aflitos et al., 2014). Furthermore, disease and pest resistance genes in tomato are primarily obtained via introgression from tomato wild relatives (Foolad et al., 2008; Moyle 2008; Ashrafi et al., 2009). Different resistance genes in each of the inbred line parents may also be contributing to heterosis via the interaction of such genes and genetic regions from other tomato species.

The interaction between the recurrent parent genetic background, E6203, of the IL parents and the male inbred lines may partly explain the presence of heterosis (Wang et al., 2009). Of the three male inbred line parents, CTRI 1558 is the most genetically distinct from other California processing tomato germplasm including E6203 that were evaluated by Park et al. (2004). The observed heterosis in certain IL x CTRI 1558 hybrids (Figure 1) may be a result of the genetic distance between CTRI 1558 and E6203, plus the presence of an introgression from *S. habrochaites*.

Heterosis has been observed in specific hybrid combinations for yield-related traits in tomato (Monforte and Tanksley 2000(b); Krieger et al., 2010; Liu et al., 2021). Heterosis for yield has been identified in a hybrid derived from *S. habrochaites,* TA532 (Monforte and Tanksley 2000(b)). TA532 is derived from the same *S. habrochaites* (LA1777) accession as the ILs

used in this study. Furthermore, LA3913 and TA532 contain an *S. habrochaites* introgression in the same region of chromosome 1. The findings by Monforte and Tanksley (2000(b)) provides further evidence that the *S. habrochaites* introgression in LA3913 contributes to heterosis in its hybrid offspring for yield-related traits.

Irrigation Treatment and Water Use Reduction

Existing horticultural practices for processing tomato production in California include the reduction of irrigation at the end of season (Hartz et al., 2008; Ayars et al., 2015). End of season irrigation reduction facilitates harvest machinery use for once over (destructive) harvest, to increase fruit Brix content, and to decrease fungal disease pressure on ripe fruit prior to harvest (Hartz et al., 2008; Ayars et al., 2015). A milder deficit water treatment than the season-long severe 33% ETc deficit irrigation used in Lounsbery et al. (2016), Arms et al. (2016), or Groh et al. (2022) was selected for this experiment because it may be feasible for growers to adopt to save irrigation water while maintaining yields. Implementing a moderate deficit irrigation scheme post-fruit set at mid- to late-season is more likely to be used by growers than a season-long severe deficit irrigation scheme.

The reduced (deficit) water treatment used in our study saved water compared to the full water treatment at each location. At locations 1 and 2, both trialed in 2018, approximately 27% less water was applied using the deficit treatment compared to the full water treatment. In 2019, 57% less water was applied compared to the full water treatment. The moderate deficit water treatment used in this study, combined with tomato varieties that are bred to have improved tolerance to water-stress, could be an effective way for tomato growers to reduce

the total amount of irrigation water needed for tomato production. The amount of precipitation in California is projected to decrease as climate change continues (Mote et al., 2005; Diffenbaugh et al., 2015; Pathak et al., 2018; Ray et al., 2020; Mahoney et al., 2021). The reduction in water resources in California will limit the longer-term viability of processing tomato production in the state, suggesting that reducing irrigation water use is a prudent strategy (Lee et al., 2010; Diffenbaugh et al., 2015; Pathak et al., 2018; Ray et al., 2020).

Significant main effect interactions were detected between genotype and water treatment for RYLD1, RYLD2, and TYLD1 (Supplemental Table 1). For these three traits-location combinations, E6203 did not consistently have higher horticultural performance under the full water treatment than for the reduced water treatment (Supplemental Table 2). In general, there was no distinct pattern for RYLD1F, RYLD1R, RYLD2F, RYLD2R, TYLD1F, TYLD1R of genotypes consistently having a higher horticultural performance under the full water treatment than under reduced water treatment.

The deficit water treatment used in this study was not severe enough to induce statistically significant differences in genotype means between the full and reduced water treatments (Supplemental Table 1). Additionally, the low number of locations and their relatively similar environment could also account for the lack of statistically significant differences among genotype means between the full and reduced water treatments. Interestingly, for tomato inbred cultivars E6203, Heinz 8761, CTRI1558, and Apex 1000 that are not known to be adapted to water-stress tolerance, their genotype means for RYLD1R, RYLD2R, and TYLD1R were in some cases higher under reduced irrigation than for the full water

treatment (Supplemental Table 2). However, it is unlikely that this particular deficit irrigation treatment would be useful in the breeding of water-stress tolerance-related traits in processing tomato because it doesn't consistently induce differences in trait genotypic means between treatments. More genotypes would need to be evaluated using this mild deficit water treatment to determine if it has utility for the breeding of water-stress tolerance-related traits. Additionally, the number of locations, repetitions, and plot size could be used to account for environmental variation.

Conclusions

Climate change is limiting the amount of irrigation water available for tomato production in California (Lee et al., 2010; Pathak et al., 2018; Ray et al., 2020). A combination of genetic material improvements via breeding and modifications to standard horticultural field practices may show the most promise in reducing the amount of irrigation water necessary for processing tomato production in California. Among the IL parents, LA3933 with a chromosome 4 introgression from *S. habrochaites* has the best potential as a parent in a hybrid cultivar breeding program in processing tomato. The other four IL parents exhibited one or more horticulturally undesirable GCA estimates for TYLD or RYLD, so are unsuitable due to the importance of yield (Hartz et al., 2008; Barrios-Masias and Jackson, 2014). The water treatment used in this study is less severe than that used in previous studies on water-stress tolerance by the St. Clair lab (Arms et al., 2016; Lounsbery et al., 2016; Groh et al., 2022). Adopting a mild to moderate deficit irrigation scheme post-full fruit set, mid to late season, is more likely to be adopted by growers to save water than a season-long severe deficit irrigation scheme. The

reduced irrigation treatment used in this study (40% of ETc post-full fruit set) did not decrease yield in material that is not known to be water-stress tolerant and also provided water savings over the season.

Supplemental Table Legends

Supplemental Table 1: Summary of F test values for ANOVAs performed on trait data for the set of 15 tomato F1 hybrids, 5 IL parents, 3 tomato inbred line parents, and recurrent parent control E6203 evaluated in the field under two water treatments (full and reduced). A dash (–) indicates not included in the model, and a blank space in the field location(s) or water treatment column indicates that data for locations or water treatments were combined.

Supplemental Table 2: Trait means and Tukey mean separation groups for 15 tomato F1 hybrids, 5 IL parents, 3 tomato inbred line parents, and recurrent parent control E6203. Genotypes were evaluated in the field under two water treatments (full and reduced). Genotypes are listed to the left with male and female parents listed for each hybrid. Under each trait, trait means for each genotype are listed to the left. If a transformation was used, the untransformed trait mean is listed in parenthesis next to the transformed trait mean. Tukey group letters are listed to the right of each trait.

Supplemental Table 3: GCA estimates for 5 IL female parents and 3 tomato inbred line male parents. Below each trait, the trait means and Tukey group letters are displayed. If a transformation was used, the untransformed trait mean is listed in parenthesis next to the transformed trait mean. GCA estimates for each parent are listed for each trait.

Supplemental Table 4: SCA estimates for 15 tomato F1 hybrids. Hybrids and their parents are listed to the left. Beneath each trait, the trait means and Tukey group letters are listed. If a transformation was used, the untransformed trait mean is listed in parenthesis next to the transformed trait mean. SCA estimates are listed beside the trait mean and Tukey group letters for each hybrid.

Supplemental Table 5: Heritability estimates, broad and narrow, were calculated for each trait. Estimates for additive variance from the male and female parent, dominance variance, and error variance are listed for each trait. The ratio of h^2/H was calculated and listed for each trait. An h^2/H ratio of 1 indicates that additive variance is responsible for all genetic variance for the trait.

Acknowledgements

This research was supported by the US Department of Agriculture National Institute of Food and Agriculture competitive grants. We thank the St. Clair lab members for contributions to field and lab data collection. We also thank the UC Davis Plant Sciences field facility crew for assistance with field experiments.

References

Aflitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., Finkers, R., Wang, J., Zhang, G.Y., Li, N., Mao, L.K., Bakker, F., Dirks, R., Breit, T., Gravendeel, B., Huits, H., Struss, D., Swanson-Wagner, R., van Leeuwen, H., van Ham, R.C.H.J., Fito, L., Guignier, L., Sevilla, M., Ellul, P., Ganko, E., Kapur, A., Reclus, E., Geus, B., van de Geest, H., te Lintel Hekkert, B., van Haarst, J., Smits, L., Koops, A., Sanchez-Perez, G., van Heusden, A.W., Visser, R., Quan, Z.W., Min, J.M., Liao, L., Wang, X.L., Wang, G.B., Yue, Z., Yang, X.H., Xu, N., Schranz, E., Smets, E., Vos, R., Rauwerda, J., Ursem, R., Schuit, C., Kerns, M., van den Berg, J., Vriezen, W., Janssen, A., Datema, E., Jahrman, T., Mo-quet, F., Bonnet, J. and Peters, S. 2014. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant Journal*. 80(1): 136-148. https://doi.org/10.1111/tpj.12616

Anthon, G.E., LeStrange, M. and Barrett D.M. 2011. Changes in pH, acids, sugars and other quality parameter during extended vine holding of ripe processing tomatoes. *Journal of the Science of Food and Agriculture*. 91: 1175-1181. https://doi.org/10.1002/jsfa.4312

Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2015. High-resolution mapping of a major effect QTL from wild tomato *Solanum habrochaites* that influences water relations under root chilling. *Theoretical and Applied Genetics*. 128: 1713-1724. https://doi.org/10.1007/s00122-015-2540-y

Ashrafi, H., Kinkade, M. and Foolad, M.R. 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(11): 935-956. https://doi.org/10.1139/G09-065

Ayars, J.E., Fulton, A. and Taylor, B. 2015. Subsurface drip irrigation in California – here to stay?. *Agricultural Water Management*. 157: 39-47. https://doi.org/10.1016/j.agwat.2015.01.001

Badu-Apraku, B., Oyekunle, M., Fakorede, M.A.B., Vroh, I., Akinwale, R.O. and Aderounmu, M. 2013. Combining ability, heterotic patterns and genetic diversity of extra-early yellow inbreds under contrasting environments. *Euphytica*. 192(3): 413-433. https://doi.org/10.1007/s10681-013-0876-4

Bai, Y., Kissoudis, C., Yan, Z., Visser, R.G.F. and Van der Linden, G. 2018. Plant behaviour under combined stress: tomato responses to combined salinity and pathogen stress. *The Plant Journal*. 93: 781-793. https://doi.org/10.1111/tpj.13800

Barrett, D.M., Weakley, C., Diaz, J.V. and Watnik, M. 2007. Qualitative and nutritional differences in processing tomatoes grown under commercial organic and conventional production systems. *Journal of Food Science*. 72(9): 441-451. https://doi.org/10.1111/j.1750-3841.2007.00500.x

Barrios-Masias, F.H. and Jackson, L.E. 2014. California processing tomatoes: morphological, physiological and phenological traits associated with crop improvement during the last 80 years. *European Journal of Agronomy*. 53: 45-55. https://doi.org/10.1016/j.eja.2013.11.007

Bates, D., Maechler, M., Bolker, B., and Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1): 1-48. https://doi.org/10.48550/arXiv.1406.5823

Bauchet, G. and Causse, M. 2012. Genetic diversity in tomato (*Solanum lycopersicum*) and its wild relatives. In M. Caliskan (Ed.). *Genetic Diversity in Plants*. Pages: 133-162. Rijeka, Croatia: Intech Europe

Bernacchi, D., Beck-Bunn, T., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley S.D. 1998. Advanced backross QTL analysis in tomato. I. identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theoretical and Applied Genetics*. 97: 381-397. https://doi.org/10.1007/s001220050908

Beyene, Y., Semagn, K., Crossa, J., Mugo, S., Stlin, G.N., Tarekegne, A., Meisel, B., Sehabiague, P., Vivek, B.S. and Oikeh, S. 2016. Improving maize grain yield under drought stress and non-stress environments in sub-Saharan Africa using marker-assisted recurrent selection. *Crop Science*. 56(1): 344-353. https://doi.org/10.2135/cropsci2015.02.0135

Beyene, Y., Gowda, M., Suresh, L.M., Mugo, S., Olsen, M., Oikeh, S.O., Juma, C., Tarekegne, A. and Prasanna, B.M. 2017. Genetic analysis of tropical maize inbred lines for resistance to maize lethal necrosis disease. *Euphytica*. 213(9): 224. https://doi.org/10.1007/s10681-017-2012-3

Bhatt, R.P., Biswas, V.R. and Kumar, N. 2001. Heterosis, combining ability and genetics for vitamin C, total soluble solids and yield in tomato (*Lycopersicon esculentum*) at 1700m altitude. *Journal of Agricultural Science*. 137: 71-75. https://doi.org/10.1017/S0021859601008838

Canady, M.A., Meglic, V. and Chetelat, R.T. 2005. A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome*. 48: 685-697. https://doi.org/10.1139/g05-032

Chetelat, R.T., and Ji, Y. 2007. Cytogenetics and evolution. In *Genetic Improvement of Solanaceous Crops* (Razdan, M.K. and Mattoo, A.K., eds). Science Publishers. Enfield, NH. Pp 77-112.

Chetelat, R.T., Pertuze, R.A., Faundez, L., Graham, E.B. and Jones, C.M. 2009. Distribution, ecology, and reproductive biology of wild tomatoes and related nightshades from the Atacama Desert region of northern Chile. *Euphytica*, 167: 77-93. https://doi.org/10.1007/s10681-008-9863-6

Chetelat, R.T., Qin, X., Tan, M., Burkart-Waco, D., Moritama, Y., Huo, X., Wills, T. and Pertuzé R. 2019. Introgression lines of *Solanum sitiens*, a wild nightshade of the Atacama Desert, in the genome of cultivated tomato. *The Plant Journal*. 100(4): 836-850. https://doi.org/10.1111/tpj.14460

Comstock, R.E., and Robinson, H.F. 1948. The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. *Biometrics,* 4: 254-266. https://doi.org/10.2307/3001412

Corrado, G., Piffanelii, P., Caramente, M., Coppola, M. and Rao, R. 2013. SNP genotyping reveals genetic diversity between cultivated landraces and contemporary varieties of tomato. *BMC Genomics*. 14: 835. https://doi.org/10.1186/1471-2164-14-835

Covarrubias-Pazaran, G. 2016. Genome assisted prediction of quantitative traits using the R package sommer. *PloS ONE*, 11(6): e0156744. https://doi.org/10/1371/journal.pone.0156744

Covarrubias-Pazaran, G. 2022. Quantitative genetics using the sommer package. *R Foun. Stat. Comput., Vienna.* https://cran.r-project.org/web/packages/sommer/vignettes/sommer.pdf (accessed 6 June. 2022)

Dariva, F.D., Copati, M.G.F., Pessoa, H.P., Alves, F.M., de Oliveira Dias, F., de Toledo Picoli, E.A., da Cunha, F.F. and Nick, C. 2020. Evaluation of anatomical and physiological traits of *Solanum pennellii* Cor. Associated with plant yield in tomato plants under water-limited conditions. *Scientific reports*, *10*(1): 1-13. https://doi.org/10.1038/s41598-020-73004-4

Das, A., Parihar, A.K., Saxena, D., Singh, D., Singha, K.D., Kushwaha, K.P.S., Chand, R., Bal, R.S., Chandra, S. and Gupta, S. 2019. Deciphering genotype-by-environment interaction for targeting test environments and rust resistant genotypes in field pea (*Pisum sativum* L.). *Frontiers in Plant Science*. 10: 825. https://doi.org/10.3389/fpls.2019.00825

Deitch, M.J., Sapundjieff, M.J. and Feirer, S.T. 2017. Characterizing precipitation variability and trends in world's Medditerranean-climate areas. *Water*. 9(4): 1-21. https://doi.org/10.3390/w9040259

De Vienne, D. and Fievet, J.B. 2020. The pitfalls of heterosis coefficients. *Plants.* 9(7): 875. https://doi.org/10.3390/plants9070875

Diffenbaugh, N.S., Swain, D.L. and Touma D. 2015. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*. 112(13): 3931-3936. https://doi.org/10.1073/pnas.1422385112

Eshed, Y. and Zamir, D. 1994. A genomic library of *Lycopersicon* pennelii in *L. esculentum*: a tool for fine mapping of genes. *Euphytica*. 79: 175-180. https://doi.org/10.1007/BF00022516

Eshed, Y., Gera, G. and Zamir, D. 1996. A genome-wide search for wild-species alleles that increase horticultural yield of processing tomatoes. *Theoretical and Applied Genetics*. 93: 877-886. https://doi.org/10.1007/BF00224089

Fehr, W.R. 1987. *Principles of Cultivar Development, vol. 1: Theory and Technique*. Iowa State University Press. Ames, Iowa. Pages 428-434

Figueiredo, A.S.T., de Resende, J.T.V., Faria, M.V., de Paula, J.T., Schwarz, K. and Zanin, D.S. 2015. Combining ability and heterosis of relevant fruit traits of tomato genotypes for industrial processing. *Crop Breeding and Applied Biotechnology*. 15: 154-161. https://doi.org/10.1590/1984-70332015v15n3a27

Figueiredo, A.S.T., Resende, J.T.V., Faria, M.V., Paula, J.T., Rizzardi, D.A. and Meert, L. 2016. Agronomic evaluation and combining ability of tomato inbred lines selected for the industrial segment. *Horticultura Brasileira*. 34: 086-092. https://doi.org/10.1590/S0102-05362016000010013

Food and Agriculture Organization of the United Nations. 2021. *FAOSTAT Crops*. FAOSTAT, Rome, Italy. http://www.fao.org/faostat/en/#data/QC (accessed 5 May. 2022).

Foolad, M.R., Merk, H.L., and Ashrafi, H. 2008. Genetics, genomics and breeding of late blight and early blight resistance in tomato. *Critical Reviews in Plant Sciences*. 27(2): 75-107. https://doi.org/10.1080/07352680802147353

Fulton, T.M., Beck-Bunn, T., Emmatty, D., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley, S.D. 1997. QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *Theoretical and Applied Genetics*. 95: 881-894. https://doi.org/10.1007/s001220050639

Fulton, T.M., Grandillo, S., Beck-Bunn, T., Fridman, E., Frampton, A., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley, S.D. 2000. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *Lycopersicon parviflorum* cross. *Theoretical and Applied Genetics*. 100: 1025-1042. https://doi.org/10.1007/s001220051384

Kulus, D. 2018. Genetic resources and selected conservation methods of tomato. *Journal of Applied Botany and Food Quality.* 91: 135-144. https://doi.org/10.5073/JABFQ/2018.091.019

Geidel, H., Weber, W.E., Mechelke, W. and Haufe, W. 2000. Selection for sugar yield in sugar beet, *Beta vulgaris*, using different selection indices. *Plant Breeding*. 119(2): 188-190. https://doi.org/10.1046/j.1439-0523.2000.00476.x

Ginkel, M.V. and Ortiz, R. 2018. Cross the best with the best, and select the best: HELP in breeding selfing crops. *Crop Science*. 58: 17-30. https://doi.org/10.2135/cropsci2017.05.0270

Glas, J.J., Schimmel, B.C.J., Alba, J.M., Escobar-Bravo, R., Schurrink, R.C. and Kant, M.R. 2012. Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences*. 13(12): 17077-17103. https://doi.org/10.3390/ijms131217077

Godoy, J.G.V. and Tesso, T.T. 2013. Analysis of juice yield, sugar content, and biomass accumulation in sorghum. *Crop Science*. 53(4): 1288-1297. https://doi.org/10.2135/cropsci2012.04.0217

Goodstal F.J., Kohler G.R., Randall L.B., Bloom A.J. and St. Clair D.A. 2005. A major QTL introgressed from wild *Lycopersicon hirsutum* confers chilling tolerance to cultivated tomato (*Lycopersicon escultentum*). *Theoretical and Applied Genetics*, 111: 898-905. https://doi.org/10.1007/s00122-005-0015-2

Grandillo, S. and Tanksley, S.D. 1996. QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species *Lycopersion pimpinellifolium*. *Theoretical and Applied Genetics*. 92: 935-951. https://doi.org/10.1007/BF00224033

Griffing, B. 1990. Use of controlled-nutrient experiment to test heterosis hypotheses. *Genetics*. 126(3): 753-767. https://doi.org/10.1093/genetics/126.3.753

Groh, A.M., Kubond, B.A., and St. Clair, D.A. 2022. Fine mapping of QTL for water use efficiency-related traits on chromosome 9 of *Solanum habrochaites* in the field. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20828

Haggard, J.E. and St. Clair, D.A. 2014. Combining ability for *Phytophthora infestans* quantitative resistance from wild tomato. *Crop Science*. 55: 240-254. https://doi.org/10.2135/cropsci2014.04.0286

Hallauer, A.R. and Miranda, J.B. 1981. *Quantitative Genetics in Maize Breeding.* Iowa State University Press. Ames, Iowa. Pages 415-416

Hartman, J.B. and St Clair, D.A. 1999. Combining ability for beet armyworm, *Spodoptera exigua*, resistance and horticultural traits of selected *Lycopersicon pennellii*-derived inbred backcross lines of tomato. *Plant Breeding*. 118: 523-530. https://doi.org/10.1046/j.1439-0523.1999.00437.x

Hartz, T.K., Miyao, G., Mickler., Lestrange, M., Stoddard., Nuñez, J., and Aegerter B. 2008. Processing Tomato Production in California. *University of California Agriculture and Natural Resources*, doi: 10.3733/ucanr.7228

Hirut, B., Shimelis, H., Fentahun, M., Bonierbale, M., Gastelo, M. and Asfaw, A. 2017. Combining ability of highland tropic adapted potato for tuber yield and yield components under drought. *PLOS One*. 12(7): e0181541. https://doi.org/10.1371/journal.pone.0181541.

Jahufer, M.Z.Z. and Casler, M.D. 2015. Application of the Smith-Hazel selection index for improving biomass yield and quality of switchgrass. *Crop Science*. 55(3): 1212-1222. https://doi.org/10.2135/cropsci2014.08.0575

Janick, J. 1998. Hybrids in horticultural crops. In K.R. Lamkey, and J.E. Staub (Ed.). *Concepts and Breeding of Heterosis in Crop Plants*. Pages: 45-56. Madison, Wisconsin, USA: CSSA Special Publication 25. https://doi.org/10.2135/cssaspecpub25.c4

Kearsey, M.J. and Pooni, H.S. 1996. The genetical analysis of quantitative traits. Chapman and Hall. New York, New York.

Krieger, U., Lippman, Z.B. and Zamir, D. 2010. The flowering gene single flower truss drives heterosis for yield in tomato. *Nature Genetics.* 42(5): 459-U138. https://doi.org/10.1038/ng.550

Kubond, B.A. and St. Clair, D.A. 2022. Bin mapping of water stress tolerance-related, fruit quality, and horticultural traits in tomato introgression lines derived from wild *Solanum habrochaites*. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20869

Lee, J., De Gryze, S., Six, J. 2010. Effect of climate change on field crop production in California's Central Valley. *Climate Change*. 109: 335-353. https://doi.org/10.1007/s10584-011-0305-4

Lenth, R.V. 2016. Least-Squares Means: The R package lsmeans. *Journal of Statistical Software*, 69(1): 1-33. https://doi.org/10.18637/jss.v069.i01

Liu, Z., Jiang, J., Ren, A., Xiangyang, X., Zhange, H., Zhao, T., Jiang, X., Sun, Y., Li, J. and Yang, H. 2021. Heterosis and combining ability analysis for fruit yield, early maturity and quality in tomato. *Agronomy*. 11(4): 807. https://doi.org/10.3390/agronomy11040807

Longin, C.F.H., Muhleisen, J., Maurer, H.P., Zhang, H., Gowda, M. and Reif, J.C. 2012. Hybrid breeding in autogamous cereals. *Theoretical and Applied Genetics*. 125: 1087-1096. https://doi.org/10.1007/s00122-012-1967-7

Lounsbery, J.K., Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2016. Quantitative Trait Loci for water-stress tolerance traits localize on chromosome 9 of wild tomato. *Crop Science*, 56: 1514-1525. https://doi.org/10.2135/cropsci2015.07.0432

Mahoney, K., Scott, J.D., Alexander, M., McCray, R., Hughes, M., Swales, D. and Bukovsky, M. 2021. Cool season precipitation projections for California and the Western United States in NA-CORDEX models. *Climate Dynamics*. 56: 3081-3102. https://doi.org/10.1007/s00382-021-05632-z

Makanda, I., Tongoona, P., Derera, J., Sibiya, J. and Fato, P. 2010. Combining ability and cultivar superiority of sorghum germplasm for grain yield across tropical low- and mid-altitude environments. *Field Crops Research*. 116(1-2): 75-85. https://doi.org/10.1016/j.fcr.2009.11.015

Makumbi, D., Betran, J.F., Banzinger, M. and Ribaut, J.M. 2011. Combining ability, heterosis and genetic diversity in tropical maize (*Zea mays* L.) under stress and non-stress conditions. *Euphytica*. 180: 143-162. https://doi.org/10.1007/s10681-010-0334-5

Marulanda, J.J., Mi, X.F., Utz, H.F., Melchinger, A.E., Wurschum, T. and Longin, C.F.H. 2021. Optimum breeding strategies using genomic and phenotypic selection for the simultaneous improvement of two traits. *Theoretical and Applied Genetics*. 134(12): 4025-4042. https://doi.org/10.1007/s00122-021-03945-5

Miller, J.C. and Tanksley, S.D. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics*. 80: 437-448. https://doi.org/10.1007/BF00226743

Monforte, A.J., and Tanksley, S.D. 2000(a). Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicum hirsutum* genome in a *L. esculentum* genetic background: a tool for gene mapping and gene discovery. *Genome*, 43(5): 803-813. https://doi.org/10.1139/g00-043

Monforte, A.J. and Tanksley, S.D., 2000(b). Fine mapping of a quantitative trait locus (QTL) from *Lycopersicon hirsutum* chromosome 1 affecting fruit characteristics and agronomic traits: breaking linkage among QTLs affecting different traits and dissection of heterosis for yield. *Theoretical and Applied Genetics.* 100:471-479. https://doi.org/10.1007/s001220050061

Mote, P.W., Hamlet, A.F., Clark, M.P. and Lettenmaier, D.P. 2005. Declining mountain snowpack in western North America. *Bulletin of the American Meterological Society*. 86(1):39-49. https://doi.org/10.1175/BAMS-86-1-39

Moyle, L.C. 2008. Ecological and evolutionary genomics in the wild tomatoes (Solanum sect. Lycopersicon). *Evolution*. 62(12): 2995-3013. https://doi.org/10.1111/j.1558-5646.2008.00487.x

Moyle L.C. and Muir C.D. 2010. Reciprocal insights into adaptation from agricultural and evolutionary studies in tomato. *Evolutionary Applications*. 3(5-6): 409-421. https://doi.org/10.1111/j.1752-4571.2010.00143.x

National Agricultural Statistics Service. 2021. *Statistics by subject*. USDA, Washington DC. http://www.nass.usda.gov/Statistics_by_Subject/?sector=CROPS (accessed 5 May 2022).

Pathak, T., Maksey, M.L., Dahlberg, J.A., Kearns, F., Balie, K.M. and Zaccaria, D. 2018. Climate change trends and impacts on California agriculture: a detailed review. *Agronomy*. 8(25): 1-27. https://doi.org/10.3390/agronomy8030025

Park, Y.H., West, M.A.L. and St. Clair, D.A. 2004. Evaluation of AFLPs for germplasm fingerprinting and assessment of genetic diversity in cultivars of tomato (*Lycopersicon esculentum* L.). *Genome*, 47(3): 510-518. https://doi.org/10.1139/g04-004

Payero, J.O., Tarkalson, D.D., Irmak, S., Davison, D. and Petersen, J.L. 2009. Effect of timing of a deficit-irrigation allocation on corn evapotranspiration, yield, water use efficiency and dry mass. *Agricultural Water Management*. 96: 1387-1397. https://doi.org/10.1016/j.agwat.2009.03.022

Processing Tomato Advisory Board. 2021. *2021 Top 50 Variety Report*. PTAB, Davis, CA. http://www.ptab.org/2021Top50.pdf (accessed 7 September 2022)

Qi, H., Huang, J., Zheng, Q., Huang, Y., Shao, R., Zhu, L., Zhang, Z., Qiu, F., Zhou, G., Zheng, Y. and Yue, B. 2013. Identification of combining ability loci for five yield-related traits in maize using a set of testcross with introgression lines. *Theoretical and Applied Genetics*. 126: 369-377. https://doi.org/10.1007/s00122-012-1985-5

Ray, P., Wi, S., Schwarz, A., Correa, M., He, M.X. and Brown, C. 2020. Vulnerability and risk: climate change and water supply from California's central valley water system. *Climate Change*. 161(1): 177-199. https://doi.org/10.1007/s10584-020-02655-z

R Core Team 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rehman, U.A., Dang, T., Qamar, S., Ilyas, A., Fatema, R., Kafle, M., Hussain, Z., Masood, S., Iqbal, S. and Shahzad, K. 2021. Revisiting plant heterosis – from field scale to molecules. *Genes*. 12(11): 1688. https://doi.org/10.3390/genes12111688

Rice, R.R. and Tracy, W.F. 2013. Combining ability and acceptability of temperate sweet corn inbreds derived from exotic germplasm. *Journal of the American Society of Horticultural Sciences.* 138(6): 461-469. https://doi.org/10.21273/JASHS.138.6.461

Richards, R.A. 2006. Physiological traits used in the breeding of new cultivars for water-scare environments. *Agricultural Water Management*. 80: 197-211. https://doi.org/10.1016/j.agwat.2005.07.013

Rick C.M. 1973. Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. *Genes, Enzymes, and Populations*. 255-269. https://doi.org/10.1007/978-1-4684-2880-3_17

Rick C.M. 1983. Genetic variability in tomato species. *Plant Mol. Biol. Rep.* 1: 81-87. https://doi.org/10.1007/BF02680303

Rothan, C., Diouf, I. and Causse, M. 2019. Trait discovery and editing in tomato. *Plant Journal*. 97(1): 73-90. https://doi.org/10.1111/tpj.14152

Ruan, Y.L., Patrick, J.W., Bouzayen, M., Osorio, S. and Fernie, A.R. 2012. Molecular regulation of seed and fruit set. *Trends in Plant Science.* 17(11): 656-665. https://doi.org/10.1016/j.tplants.2012.06.005

Schouten, H.J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y. and Visser, R.G.F. 2019. Breeding has increased the diversity of cultivated tomato in The Netherlands. *Frontiers in Plant Science*. 10: 1606. https://doi.org/10.3389/fpls.2019.01606

Seager, R., Osborn, T.J., Kushnir, Y., Simpson, I.R., Nakamura, J. and Liu, H.B. 2019. Climate variability and change of Mediterranean-type climates. *Journal of Climate*. 32(10): 2887-2915. https://doi.org/10.1175/JCLI-D-18-0472.1

Spooner, D.M., Peralta, E. and Knapp, S. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes (*Solanum L.* section *Lycopersicon* (Mill.) Wettst.). *Taxon*. 54:43-61.47. https://doi.org/10.2307/25065301

Sprague, G.F. and Tatum, L.A. 1942. General vs. specific combining ability in single crosses of corn. *Journal of the American Society of Agronomy*. 34: 923-932. https://doi.org/10.2134/agronj1942.00021962003400100008x

Tadesse, W., Morgounov, A.I., Braun, H.J., Akin, B., Keser, M., Kaya, Y., Sharma, R.C., Rajaram, S., Singh, M., Baum, M. and Ginkel, M.V. 2013. Breeding progress for yield in winter wheat genotypes targeted to irrigated environments of the CWANA region. *Euphytica*. 194: 177-185. https://doi.org/10.1007/s10681-013-0903-5

Tamburino, R., Sannino, L., Cafasso, D., Cantarella, C., Orru, L., Cardi, T., Cozzolino, S., D'Agostino, N. and Scotti, N. 2020. Cultivated tomato (*Solanum lycopersicum* L.) suffered a severe cytoplasmic bottleneck during domestication: implications from chloroplast genomes. *Plants-Basel*. 9(11): . https://doi.org/10.3390/plants9111443

Tanksley, S.D., Grandillo, S., Fulton, T.M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. and Beck-Bunn T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *S. pimpinellifolium*. *Theoretical and Applied Genetics*. 92: 213-224. https://doi.org/10.1007/BF00223378

Truco, M.J., Randall, L.B., Bloom A.J. and St. Clair D.A. 2000. Detection of QTLs associated with shoot wilting and root ammonium uptake under shilling temperatures in an interspecific backcross population from *Lycopersicum esculentum* x *L. hirsutum*. *Theoretical and Applied Genetics*. 101:1092-1092. https://doi.org/10.1007/s001220051583

Tuberosa, R. 2012. Phenotyping for drought tolerance of crops in the genomic era. *Frontiers in Physiology*. 3(347): 1-26. https://doi.org/10.3389/fphys.2012.00347

Wang G.L., Kang, M.S. and Moreno, O. 1999. Genetic analyses of grain-filling rate and duration in maize. *Field Crops Research*. 61(3): 211-222. https://doi.org/10.1016/S0378-4290(98)00163-4

Wang, L., Zhang, Z., Wei, L., Zhang, D., Teng, F., Tao., Y. and Zheng, Y. 2009. The residual background genome from a donor within an improved line selected by marker-assisted selection: impact on phenotype and combining ability. *Plant Breeding*. 128: 429-435. https://doi.org/10.1111/j.1439-0523.2008.01611.x

Wehner, T.C. 1999. Heterosis in vegetable crops. *Genetics and Exploitation of Heterosis in Crops*. In J.G. Coors and S. Pandey (Ed.). *Genetics and Exploitation of Heterosis in Crops*. Pages: 387-397. Madison, Wisconsin, USA: American Society of Agronomy, Crop Science Society of America

Williams, J.S. 1962. The evaluation of a selection index. *Biometrics*. 18(3): 375-393. https://doi.org/10.2307/2527479

Yang, X.H., Caro, M., Hutton, S.F., Scott, J.W., Guo, Y.M., Wang, X.X., Rashid, M.H., Szinay, D., de Jong, H., Visser, R.G.F., Bai, Y.L. and Du, Y.C. 2014. Fine mapping of the tomato yellow leaf curl virus resistance gene Ty-2 on chromosome 11 of tomato. *Molecular Breeding.* 34(2): 749-760. https://doi.org/10.1007/s11032-014-0072-9

Zhang, J.F., Wu, M., Yu, J.W., Li, X.L. and Pei, W.F. 2016. Breeding potential of introgression lines developed from interspecific crossing between upland cotton (*Gossypium hirsutum*) and *Gossypium barbadense*: heterosis, combining ability, and genetic effects. *PLOS One*. 11(1): e0143646. https://doi.org/10.1371/journal.pone.0143646

Chapter 3:

Evaluation of a Tractor-Based High-Throughput Field Phenotyping System for Potential Use in Tomato Breeding

Abstract:

Global climate change is reducing the amount of precipitation that provides fresh water for crop production in arid regions. Cultivated tomato, *Solanum lycopersicum*, is susceptible to water-stress, which reduces yields. In order to enhance breeding efficiency, large amounts of accurate phenotype data must be readily obtainable and available to complement quick and inexpensive genotype data. High-throughput phenotyping (HTP) techniques are being developed and tested to address the lack of phenotype data. Here, a tractor-based HTP multi-spectral imaging robot (HTP robot) was evaluated in the field using two sets of tomato genotypes under two water treatments (normal and reduced) to evaluate the HTP robot's potential for use in assessing ripe yield, a trait of importance to tomato breeding. The HTP robot collected RGB images that were subsequently processed to obtain "red pixel number" data, the number of red pixels present in a given subplot, and ground truthing was done by manually collecting ripe tomato yield data. The first set of tomato genotypes was used to create a model that associated HTP robot collected phenotype data and manually collected phenotype data and the second set was used to validate the association between the two sets of data. Processing terabytes of HTP robot collected images into a data form useful for analysis is extremely time consuming and consequently, we could only obtain a limited amount of HTP

data for comparison with manually collected yield data. Spearman correlations between the red pixel number and actual ripe tomato yield ranged from 0.62 to 0.76, indicating a reasonably strong correlation, suggesting HTP has a promising role in future phenotyping experiments. However, further evaluation of the HTP robot collected data for this trait and other traits of importance would be required to evaluate its full efficacy for use in tomato breeding.

Introduction:

Global climate change is leading to major effects on production environments of crop plants worldwide (Lee et al., 2010; Tester and Langridge, 2010). Environmental perturbations caused by climate change will impact global food production if not addressed (Tester and Langridge 2010; Furbank and Tester 2011; Cobb et al., 2013). Tomato (*Solanum lycopersicum*) is the second most widely grown vegetable crop worldwide and in the United States (FAOSTAT, 2021; USDA NASS, 2022). California is the largest producer of tomatoes in the United States. In 2021, California produced tomatoes with a total farmgate value of 1.18 billion dollars (USDA NASS, 2022). California produces over 95% of the processing tomatoes grown in the United States with 10.78 million tons worth 905 million dollars in farmgate value in 2021 (USDA NASS, 2022).

Climate change is expected to limit the feasibility of processing tomato production in California due to reduced precipitation and increasing temperatures (Lee et al., 2010; Diffenbaugh et al., 2015; Pathak et al., 2018; Ray et al., 2020). California has a Mediterranean climate characterized by hot dry summers and cool wet winters (Deitch et al., 2017; Pathak et

al., 2018; Seager et al., 2019). Due to the lack of precipitation during the summer growing season, the production of processing tomatoes in California relies on irrigation (Hartz et al., 2008). Winter snowpack in the Sierra Nevada mountains is an essential source of water for irrigation in California during the summer months. Climate change has led to a decrease in snowpack in the western United States from the 1940s to present and is expected to experience further decreases with continuing increases in temperatures (Mote et al., 2005; Mahoney et al., 2021). The decline in snowpack further limits water resources available for use in crop production, including tomatoes (Pathak et al., 2018; Ray et al., 2020).

Cultivated tomato possesses a limited genetic base due to multiple bottleneck events that occurred during domestication (Rick 1983; Miller and Tanksley, 1990; Corrado et al., 2013; Kulus, 2018; Tamburino et al., 2020). In contrast, wild species of tomato are an abundant source of genetic diversity for expanding the genetic base of cultivated tomato for crop improvement. Wild tomato relatives have been used previously in breeding as a source of valuable traits including both biotic and abiotic stress tolerances (Rick, 1983; Bai et al., 2018; Schouten et al., 2019). The genomes of wild species of tomato and cultivated tomato share a large degree of synteny and the species are inter-crossable (Spooner et al., 2005; Chetelat and Ji 2007). Previous work in the St. Clair lab has mapped quantitative trait loci (QTL) for water-stress tolerance-related traits to chromosome 9 of *S. habrochaites* (Goodstal et al., 2005; Arms et al., 2015; Lounsbery et al., 2016; Groh et al., 2022). Additionally, other tomato wild relatives, including *S. lycopersicum* var. *cerasiforme*, *S. pennellii*, *S. chilense*, and *S. sitiens* have been

reported to be tolerant to water-stress (Dariva et al., 2020; Chetelat et al., 2009; Moyle and Muir 2010; Rick 1973).

With the advent of inexpensive ways to obtain genotype data, genomic methods have been developed to increase the speed and efficiency of the plant breeding process (Bernardo 2008; Walsh 2009). One of the major limiting factors in utilizing high throughput breeding methods is the need for obtaining large amounts of phenotype data to accompany the genotype data (Bernardo 2008; Walsh 2009; Cobb et al., 2013; Araus and Cairns, 2014). Traditionally, phenotypic data has been time consuming and expensive to collect due to reliance on manual methods employed by plant breeders and collection techniques unique to each crop (Cobb et al., 2013; Araus et al., 2018). To help address the gap between the amount of available genotype and phenotype data, high-throughput phenotyping (HTP) methods are being developed and tested (Araus and Cairns 2014; Fahlgren et al., 2015). Ideally, robot collected phenotype data are less expensive and time consuming to obtain than traditional methods. Larger amounts of data can also be collected by phenotyping robots than by traditional methods. Developing HTP methods to fully utilize high throughput genotyping datasets could be useful in accelerating the plant breeding process to address crop improvement for a rapidly changing climate (Furbank and Tester, 2011).

HTP methods have primarily been developed for agronomic crops, including but not limited to corn (*Zea mays*), wheat (*Triticum aestivum*), sorghum (*Sorghum bicolor*), and rice (*Oryza sativa*) (Araus et al., 2018; Atefi et al., 2021). These crops are all grasses with similar and

relatively simple plant architecture, facilitating the application of HTP platforms. In contrast, vegetables constitute a large number of diverse species with complex plant architecture often leading to occlusion of fruit by vegetation, making accurate image analysis of ultimate traits like yield difficult (Atefi et al., 2021; Fonteijn et al., 2021). Thus, developing effective HTP systems for vegetables is more technically challenging than for many agronomic crops.

HTP sensors can be mounted on stationary platforms or on hand-held, tractor or boom-based, or aerial drone-based systems (Araus and Cairns, 2014; Yang et al., 2017; Araus et al., 2018; Kaur et al., 2020; Song et al., 2021). Each platform has different positives and negatives for use in plant breeding. The most common platform used in high-throughput phenotyping are aerial drones due to their relatively low cost and ease of use (Araus and Cairns, 2014; Yang et al., 2017). To date, only one report has been published using drones to phenotype tomatoes in the field (Chang et al., 2021). Drones are currently unable to overcome the issues caused by the complex plant architecture of tomato (Roo, 2022). HTP systems have been developed to phenotype tomatoes in controlled greenhouse settings (Halperin et al., 2016; Fonteijn et al., 2021).

Tractor-based HTP platforms have several advantages when phenotyping vegetables compared to other HTP platforms. Tractor-based phenotyping systems are mobile, allowing for phenotype data collection from many locations. Other ground-based systems are not easily transported or are completely stationary once installed (Yang et al., 2017; Song et al., 2021). Multi-location trialing is an essential part of the breeding process and is required to effectively

develop varieties suitable to a range of target environments (Cobb et al., 2013). Tractor-based platforms can be placed physically closer to plants during phenotyping compared to other types of mobile HTP platforms such as drones (Roo, 2022). Tractor-based platforms can also be fitted with supplemental lighting to enable uniform illumination and better-quality sensor data and can include a larger number of sensors than drone-based platforms (Song et al., 2021; Roo, 2022). HTP platforms developed for greenhouses and other controlled environments do not necessarily facilitate improvements for crops grown in field environments. Because plant breeding must take place in the target environments for crop production, systems developed for use in controlled environments are not often useful for accurate evaluation of field-grown crops, rendering them less effective than HTP platforms that can phenotype in the target environments (Araus and Cairns, 2014; Atefi et al., 2021).

The original objective for this experiment was to determine if relationships exist between manually collected phenotype data and data collected by a tractor-based HTP system. The aim was to determine if the HTP robot was capable of collecting images that could be processed into useable data with sufficient predictive accuracy to be potentially useful in tomato breeding. Unfortunately, due to unforeseen complications, only a limited amount of HTP data were processed into a form that could be compared to the manual tomato phenotype data. However, we were able to conduct a limited preliminary comparison of HTP data for red pixels and the corresponding manually collected ripe fruit yield data.

Materials and Methods:

Plant Material

Two sets of plant material were used (Supplemental Table 1). In 2017 and 2018, a phenotypically diverse set of tomato genotypes was evaluated including introgression lines (ILs) derived from a wild tomato relative, *S. habrochaites*, in the genetic background of *S. lycopersicum* inbred processing cv. E6203 (Monforte and Tanksley 2000) and a set of inbred processing tomato cultivars. A large range of phenotypic diversity for key processing tomato traits such as yield, maturity, and shoot size was included in this set. In 2017, one genotype, DRI 319, was determined to have been a seed mix up and was not included in the analysis. It was the replaced in 2018 by cultivar Orion. In 2019, ILs, inbred processing tomato cultivars, and modern commercial F1 hybrids and a number of genotypes used in the first two years were evaluated.

Field Experimental Design

Field experiments were conducted at the UC Davis Plant Sciences Field Research Facility in Davis, California in the summers of 2017, 2018, and 2019. Each year, trials were planted at two locations, representing fields with different cropping histories. The experimental design was a randomized complete block with a split-plot treatment design. Main plots consisted of irrigation treatments and subplots consisted of genotypes. The experiment included two replications at each location in 2017 and one replication per location in 2018 and 2019. Main-plots received one of two irrigation treatments, full crop evapotranspiration (ETc) for the

season or reduced water post-fruit set. The deficit treatment was 40% ETc after E6203 reached full fruit set, which was a 60% reduction in water applied.

In 2017 and 2018, subplots included 28 genotypes; in 2019, they included 40 genotypes. Main-plots consisted of 4 blocks of subplots. In 2019, main-plots consisted of 3 blocks of subplots. Each subplot of a given genotype consisted of 10 plants spaced 30.5cm apart in a single row in the center of a 154.2cm bed. A within-row alley of 121.9cm without plants was included between the end and the beginning of each subplot to facilitate manual and robotic data collection. Double border rows of inbred processing tomato cultivars were planted between each main-plot and at the outer perimeter of the field experiment to minimize edge effects.

Seeds were planted into flats in the greenhouse on the UC Davis campus during March each year. Plants were grown until they reached the second true leaf stage, approximately 5 weeks after seeding, when they were transferred to a lath house to harden off for about a week before hand transplanting to the field. After transplanting, seedlings were sprinkler irrigated for approximately 2-3 weeks to allow for plant establishment. Subsequently, water was applied via subsurface drip irrigation with 1.6 cm width drip tape and 30.5 cm emitter spacing (Toro Flow Control). The drip tape was buried 20 - 22cm below the soil surface of the bed prior to transplant. Water was delivered three times weekly to evenly distribute the irrigation water application over the week and to enable the field surface to dry sufficiently for the HTP tractor to enter the field for data collection. The amount of water delivered during each irrigation was

determined by the amount of evapotranspiration (ET) recorded by a nearby CIMIS station (cimis.water.ca.gov) since the previous irrigation and by the canopy size of the growing tomato plants. Because water pressure fluctuated daily, the water flow rate was calculated before each irrigation. The length of each irrigation treatment was calculated using the flow rate of the irrigation system and the calculated amount of water to apply to each treatment as determined by ETc for tomato and canopy size. The irrigation system was manually turned on and off at the beginning and end of each irrigation period. Flow meters (Sensus VMRS2 Brass Water Meter) were used to record the amount of water applied to each treatment and location. No measurable summer precipitation was recorded in 2017 and 2018. In 2019, 7.3 cm of rain was recorded in May (cimis.water.ca.gov).

Urea nitrogen fertilizer was banded prior to transplant at a depth of 15-20cm beneath the soil surface at a rate of 22.4 Kg ha^{-1} (Hartz et al., 2008). Additional urea nitrogen was applied by fertigation for four weeks at a rate of 33.6 Kg ha^{-1} starting 21 days post-transplant. A total of 156.8 Kg/ha of supplemental Urea nitrogen fertilizer was applied through the duration of the season.

Manual Phenotyping of Traits

Manual phenotype data was collected on a per-subplot basis in each year of the experiment. Traits were grouped into Horticultural and Water-Stress Tolerance-Related categories (Table 1).

Table 1: Traits evaluated by manual phenotyping and tractor-mounted HTP robot phenotyping in a diverse set of processing tomato genotypes, by category. Habit, Canopy, and Sunburn were evaluated only in 2017 and 2019.

Trait Category	Trait Code	Description
Water-Stress Tolerance-Related	SDW	Shoot dry weight (biomass, in g)
	SFW	Shoot fresh weight (biomass, in g)
	BioM	Total fresh above ground biomass at harvest (g)
Horticultural	RYLD	Ripe fruit yield (kg)
	TYLD	Total fruit yield (kg)
	DAPG	Days after planting to first green fruit
	DAPR	Days after planting to first ripe fruit
	Canopy	Leaf canopy cover (score 1-5, sparse to dense)
	Sunburn	Degree of fruit sunburn (score 1-5, none to severe)
	Habit	Plant growth habit (score 1-5, prostrate to erect)
	HI	Harvest index (ratio of ripe fruit yield to shoot fresh weight and total fruit yield)

Days after planting to first green fruit (DAPG) was defined as the number of days from transplanting until greater than 50% of the plants in a subplot each had a green fruit of at least 1 cm in diameter. Scoring was conducted on Monday, Wednesday, and Friday each week. Days to first red fruit (DAPR) was defined as the number of days from transplanting until greater than 50% of the plants in a subplot had at least one fully ripe fruit (Lounsbery et al., 2016) and was scored on the same schedule as DAPG.

Plant growth habit, degree of fruit sunburn, and canopy cover were scored when each subplot immediately prior to harvest. Subplots were subjectively scored for habit on a scale of 1 to 5, ranging from very prostrate to very upright. Sunburn and Canopy were also scored on a subjective 1 to 5 scale, from no fruit sunburn to severe sunburn, and sparse canopy cover to dense canopy cover, respectively.

Subplots were destructively harvested when all the plants in a subplot had approximately 90-95% ripe fruit. If a subplot was later maturing relative to control cv. E6203, it was harvested at the end of the growing season regardless of the ripe fruit load. Additional details on destructive harvest procedures can be found in Lounsbery et al. (2016) and Kubond and St. Clair (2022; Chapter 2). After fruit were removed and weighed, fresh tomato shoots were weighed, placed in onion mesh bags, put in a forced air dryer at 60 ^{o}C for at least 2 weeks, then weighed to obtain shoot dry weight per subplot (Lounsbery et al., 2016). Total fresh biomass was determined by adding total fruit yield to shoot fresh weight. Harvest index was calculated using the following formula: HI = RYLD/(TYLD + SFW). Trait data for TYLD, RYLD, HI, SDW, SFW, and BioM was only collected at one location in 2017 to correspond with the subplots that the HTP robot collected data on.

HTP Robot/HTP Robot Phenotyping

The HTP robot developed by Dr. David Slaughter and colleagues in the UC Davis Biological and Agricultural Engineering Department included 4 types of sensors, including standard color cameras, infrared cameras, time-of-flight sensors, and thermal sensors (Vuong

et al., 2020; Roo et al., 2020; Roo, 2022). Each of the sensors was capable of collecting data potentially relevant to a number of traits evaluated manually. The standard color cameras were used to phenotype TYLD, RYLD, DAPR; the infrared cameras, DAPG and DAPR; the time-of-flight sensors, SFW, BioM, Habit, and Canopy; and the thermal sensors, leaf temperature. Manual phenotype data was not collected for leaf temperature due to the lack of resources and labor required.

The HTP robot was manually driven through the experimental fields twice a week to collect data on all subplots row by row, starting 4 weeks post-transplant and continuing through the duration of the season. The HTP robot took approximately 5-6 hours to collect data each time it was run through the field. The robot would begin data collection typically around 10am and would operate through rows of the experiment as field conditions permitted. The same course was not run each time by the robot. The relatively long period for data collection can limit the effectiveness of the HTP robot to collect trait data that is highly dependent on environment conditions such as canopy temperature (Deery et al., 2016). Image data from each subplot were digitally merged into a three-dimensional re-creation of the subplot. Point cloud data collected from the time-of-flight sensors were also reconstructed into three-dimensional re-creations of each subplot. Image data were collected by the robot for most subplots in each year.

Ripe fruit yield was quantified using red pixel number (RPN), total number of red pixels present in the computer-generated reconstruction from visual spectrum images of each

subplot. For this chapter, we assessed RPN from one main-plot of the experiment in 2017 at two separate time points.

Statistical Analysis

Statistical analysis for the manual phenotype data was performed on a per trait basis. RPN was analyzed in parallel with the manually collected phenotype data. Normality was assessed for each trait using Shapiro-Wilk W-Statistic and Quantile-Quantile plots and homogeneity of variance was evaluated using residuals plots. Analysis of variance (ANOVA) was performed for each trait using mixed linear models. Statistical analyses were accomplished using the R statistical package (R Core Team) and the lme4 package (Bates et al., 2015).

Trait data from the two training years, 2017 and 2018, was initially analyzed across both years. Because significant year x genotype interactions were detected for each trait, ANOVAs were performed separately for each year using the following linear additive model:

Trait = Loc + Rep(loc) + Water + Genotype + (Water x Loc) + (Genotype x Loc) + (Genotype x Water) + (Genotype x Water x Loc)

Loc refers to field location (location 1 or 2) within each year. Water refers to the water treatment (Full ETc or reduced irrigation post-fruit set). Genotype refers to the set of genotypes used in a given experiment (see Supplemental Table 1). When significant ($P \leq 0.05$) interactions were detected for a particular trait, analysis was performed by year, location, treatment, or a combination of factors, depending on the source of variation exhibited in an interaction. For the analysis, water, repetition, and genotype were considered fixed effects and location was

considered random. The same linear additive model was used to evaluate phenotypic data collected in the validation year, 2019.

The following linear additive model was used to evaluate RPN and the corresponding RYLD dataset for 2017:

Trait = Genotype

After ANOVA was conducted for each trait, mean separations were performed with Tukey's test, which compared each genotype to every other genotype, with the cld(contrasts) function in the lsmeans package in R (Lenth et al., 2016). When significant G x E interactions (genotype x location or genotype x water treatment) were detected, trait data were referred to by their specific year, location, and treatment combinations. For example, total fruit yield (TYLD) from location 1 under the full water treatment in 2018 is denoted as TYLD181F, with the year being appended after the trait code, either 17 for 2017, 18 for 2018, or 19 for 2019, followed by the location (1 or 2) and finally the water treatment, F for full water treatment or R for reduced water treatment.

Comparison of Manually and HTP Robot collected Data

Manually collected RYLD was compared to RPN to determine what relationship may exist between the two traits. Linear regressions were performed on the lsmeans for RYLD and RPN at 107 and 114 days post-transplant for each genotype using Microsoft Excel. Spearman rank correlations between the three traits were computed using cor.test function in the R stats package (R Core Team).

Results

Manual Phenotype Data

ANOVAs and mean separations were performed for each manually collected trait (Supplemental Table 2; Supplemental Table 3; Supplemental Table 4; Supplemental Table 5). Three traits, Canopy, Habit, and Sunburn were collected only in 2017 and 2019. Statistically significant genotype main effects were observed for each trait except for TYLD181F and BioM181F (Supplemental Table 2; Supplemental Table 3). Mean separations were not performed for these two traits as there were no significant genotype effects detected using ANOVA.

Within the training data set, 2017 and 2018, genotypes differed for key yield and maturity traits. Trait means for TYLD ranged from 3.53 kg to 20.58 kg, RYLD from 2.52 kg to 20.18 kg, DAPG from 45 days to 75.3 days, and DAPR from 80 days to 109.6 days (Supplemental Table 4). The wide range of trait means among genotypes for each trait was desirable in order to train the HTP robot on a large range of possible tomato phenotypes. In general, higher yields as well as later maturity were observed in 2017 compared to 2018,

In contrast to the training data, the validation data set of 2019 displayed a narrower range of trait means among the genotypes (Supplemental Table 5). Trait means in the validation set for TYLD ranged from 7.45 kg to 13.52 kg, RYLD from 5.63 kg to 11.78 kg, DAPG from 30.8 days to 45 days, and DAPR from 71.5 days to 84.3 days (Supplemental Table 5). The validation set was limited to one year, possibly limiting the amount of phenotypic variation that

would be observed compared to a multi-year trial. Additionally, the material used in the 2019 validation data set was predominantly high yielding modern commercial F1 hybrid cultivars (Supplemental Table 1).

HTP Robot Phenotype Data

Image and time-of-flight data were collected by the HTP robot for approximately 16 weeks during each season. Despite the HTP robot successfully collecting image data for most of the experimental subplots, only trait data for RPN at two time points, 107 and 114 days after transplant, was available for our analysis here because processing and analysis of image data was a far more difficult and complex task than anticipated. As a result of the unexpected data processing complexity, RPN data was available for one reduced water main-plot of one location in 2017. Subsequently, ANOVAs and mean separations were performed for each of the two RPN time points (Table 3). Significant differences among genotypes were detected for both the 107 and 114 days after transplant RPN data (Supplemental Table 6).

Table 3: Trait means for the training germplasm for RPN taken at 107 and 114 days post transplant. Data was collected from a single reduced water main-plot from a single location in 2017. Genotypes are listed to the left. Under each trait, trait means for each accession are listed to the left. The standard error for each dataset is listed to the bottom of the table.

Genotype:	Chr	RPN 107	RPN 114
LA3913	1	59630	87603
LA3918	1	76042	84939
LA3921	2	39827	54863

LA3922	2	107734	240547
LA3927	3	28280	225712
LA3930	4	81843	125493
LA3933	4	76602	158863
LA3938	5	21407	43131
LA3939	5	90744	214061
LA3943	5	18497	10705
LA3948	7	14317	29500
LA3951	7	131761	290785
LA3953	8	46865	58305
LA3955	8	117766	306268
LA3956	9	86499	247378
LA3957	9	63102	217219
LA3958	9	19346	5954
LA3960	9, 10	42810	30638
LA3963	10	18915	46096
LA3965	2, 10, 11	4305	17050
LA3967	11	44212	104534
LA3968	12	48449	158436
LA3969	12	85162	260658
LA3975	3	56046	137420
E6203		72160	220933
Hunt 100		14690	76773
UC204B		59048	314597
Peto 95-43		58862	263964
Standard Error:		6295	19085

Comparison of Manual and HTP Robot Collected Data

A corresponding subset of manually collected RYLD data was compared to RPN using both Spearman rank correlations and linear regression. Significant ($P \leq 0.05$) correlations were detected between RYLD and RPN 107 days (r = 0.62) and RPN 114 days (r = 0.76) as well as between RPN 107 days and RPN 114 days (r = 0.78). Linear regressions were R^2=0.36 between RYLD and RPN 107 days and R^2=0.60 between RYLD and RPN 114 days. LA3969, LA3956, E6203, LA3955, and LA3957 had the highest RYLD of the training genotypes. Of these genotypes, LA3956 and LA3955 were among the five highest RPN values among the training genotypes for the 107 days data. LA3969 and LA3955 were also found among the top genotypes in the 114 days data.

Discussion

Limitations of HTP Systems

A number of physical limitations of the tractor-based HTP robot used in this study were identified. This prototype HTP robot was only able to enter the field to phenotype under ideal conditions. Surface moisture due to routine leaks in the drip irrigation system prevented the HTP robot from entering certain rows on some measurement days (Roo, 2022). Song et al. (2021) previously identified problems using ground based mobile HTP platforms in sub-optimal field conditions. Additionally, similar to what other researchers have found, the HTP robot was sensitive to ambient light conditions and could only phenotype plants when the sun was

sufficiently high in the sky each morning (Song et al., 2021; Roo, 2022). This decreased the length of time each day that the HTP robot could operate for data collection.

In addition to physical limitations of the HTP robot, other issues and drawbacks were identified. First, the HTP robot collected in excess of a terabyte of data during each phenotyping run and consequently, we only collected data twice weekly. Nevertheless, the sheer volume of data collected by the robot was overwhelming and processing it into a form useful for statistical analysis limited the scope of our experiment described here. As a result, only data for RPN at one main-plot at a single location in a single year was processed from the larger amount of image data collected by the HTP robot.

Similar to what was found in our study, Fonteijn et al. (2021) identified issues in the creation of digital reconstructions of tomato plants for use in phenotyping. Complex plant architecture creates challenges when evaluating HTP platforms for use in vegetable breeding (Atefi et al., 2021; Fonteijn et al., 2021). This high degree of architectural complexity leads to issues in obtaining accurate digital reconstructions of single plants, or of multiple plants in a plot that accurately represents the plots (Fonteijn et al., 2021). Previous HTP studies in tomato were performed on many fewer total number of plants or plots than used in our study. The previous greenhouse-based studies used fewer than 150 plants in individual plots in contrast to the large replicated design used in our field study (Halperin et al., 2016; Fonteijn et al., 2021). Chang et al. (2021) used approximately 100 plots with 4 plants each at a single field location in a single year. Our study had approximately 950 subplots in 2017, and approximately 500

subplots each in 2018 and 2019, each containing 10 plants. The large amount of plant material relative to previous HTP studies in tomato exacerbated issues with image data analysis, leading to the overall lack of robot-derived data to compare to the manual phenotype data.

The huge amount of data collected by HTP systems leads to complications with data processing, data management, and data analysis to obtain useful parameters (Araus and Cairns, 2014; Kar et al., 2020; Fonteijn et al., 2021; Song et al., 2021). Our study identified these same issues. The HTP robot in our study collected data onto an internal set of computer storage drives. These data had to be manually transferred from the HTP robot to other computing devices for analysis. The amount of image data collected by the HTP robot was orders of magnitude higher, 20,000 plants vs approximately 100-400 plants, than in other studies in tomato, further exacerbating known issues with image data analysis (Fonteijn et al., 2021). Challenges in analyzing large datasets are not unique to HTP and have been identified during the use of “Big Data” in other subject areas (Wu et al., 2014; Zhang et al., 2018). Solutions to the identified short comings of HTP technology in terms of data processing, management, and analysis may be identified in the larger data science literature because these limitations exist in other fields of study.

Comparison of RYLD and RPN

Manually collected RYLD data and robot collected RPN data were compared in order to determine if any relationship existed between the two data sets. Among the 5 genotypes with the highest RPN values, 2 at each time point corresponded to 2 of the 5 genotypes with the

highest RYLD values. Statistically significant Spearman rank correlations between RYLD and RPN datasets yielded values of r = 0.620 and r = 0.760 for the 107 and 114 days RPN datasets, respectively. These correlation coefficient values indicated that a moderate correlation exists between RYLD and 107 days post-transplant RPN and a strong correlation exists between RYLD and 114 days post-transplant. Linear regressions did not show as strong an association between RYLD and RPN as Spearman rank correlations found. R^2 values of 0.364 and 0.603 for the 107 and 114 days post-transplant RPN datasets, respectively, indicate that RPN can explain a moderate amount of variation for RYLD.

No studies to date have been published using red pixel number as proxy trait for yield in tomato. In wheat, red pixel count of non-vegetative segments of images taken at the wheat booting phase were found to be representative of the final grain yield (Haghshenas and Emam, 2019). Haghshenas and Emam (2019) found correlations with $R^2 = 0.690$ and correlation coefficient of 0.831 between red pixel count of non-vegetative segments of images data and final grain yield. These values are similar, though slightly higher than those identified between RYLD and the 114 days post-transplant RPN dataset. In addition to acting as a proxy for yield, red pixel number has been used in *Arabidopsis thaliana*, wheat, and soybean to collect proxy data for traits including leaf anthocyanin, trichome density, and final grain yield (Ispiryan et al., 2013; Bai et al., 2016; Haghsenas and Emam, 2019).

Two previous HTP studies in tomato have identified a relationship between yield and HTP robot collected data (Chang et al., 2021; Fonteijn et al., 2021). Chang et al. (2021) used a UAV platform to collect visual spectrum image data and multispectral data. Using color spectra

data, an excessive greenness index (ExG index) was calculated. Various features of the ExG index were used to predict total yield, with strong correlations ($R^2 > 0.7$) identified (Chang et al., 2021). Fonteijn et al. (2021) used a commercial robot (IRIS!, Metazet-Formflex) in controlled conditions in a greenhouse to collect visual spectrum image data. Image data collected prior to harvest and post-harvest by the HTP robot were compared to each other to obtain estimates for yield. A moderate correlation of r = 0.43 was established between this HTP system and manually collected yield data (Fonteijn et al., 2021). This greenhouse-based system was unable to establish as strong of correlations between HTP robot data and manually collected phenotype data as either the UAV based system used in Chang et al. (2021) or in our study.

The amount of RPN data was very limited for this experiment, unfortunately. A single main-plot from one location in one year had RPN data collected and analyzed. As a consequence, the resulting correlations obtained between RYLD and RPN are limited in their assessment of the ability for the HTP tractor to accurately collect predictive phenotype data. The phenotype for a particular trait is a result of the combination of genetic and environment as well as the interaction between genetics and the environment (Fehr 1987; Bernardo 2008; Chang et al., 2021). The limited amount of RPN data obtained limits the scope of the conclusions that can be made about the efficacy of the HTP robot for collecting data that is predictive of ripe yield. Plant breeders evaluate material over a range of target environments when making evaluations of material (Fehr 1987; Bernardo 2008; Chang et al., 2021). Further trialing of the HTP robot at multiple locations would be needed to better determine how representative the HTP robot collected RPN data is compared to the manually collected RYLD

data (Chang et al., 2021). Nevertheless, the preliminary correlations we found here are certainly promising. Additionally, the HTP robot would need to be evaluated using a different set of genetic material, such as that used in 2019 to determine its efficacy outside of the training material.

Water Use Saving using Reduced Water Treatment

The reduced water irrigation treatment contributed to a 18% to 42% water use reduction compared to the full water treatment over the length of the season depending on year and location. Current horticultural practices for processing tomato include a reduction of irrigation applied at the end of the growing season (Hartz et al., 2008; Ayars et al., 2015). The purpose of the current end of season reduction in irrigation water is to allow the entry of machinery destructive harvest, increase fruit Brix content, and to decrease fungal pressure on ripe fruit prior to harvest (Hartz et al., 2008; Ayars et al., 2015). Previous studies on water-stress tolerance in the St. Clair lab used a more severe 33% ETc season-long deficit irrigation treatment (Lounsbery et al., 2016; Arms et al., 2016; Groh et al., 2022). A less severe treatment was selected for this study due it being more feasible for growers to potentially adopt to save water if it was shown to not negatively impact yield.

The total amount of precipitation in California is expected to continue to decrease as climate change continues (Mote et al., 2005; Diffenbaugh et al., 2015; Pathak et al., 2018; Ray et al., 2020; Mahoney et al., 2021). The reduction in water resources in California is expected to limit the viability of continued processing tomato production in the state as the climate changes (Pathak et al., 2018; Ray et al., 2020). A multifaceted approach to reducing water use in

processing tomato production is needed with a rapidly changing climate. A combination of genetic material improvements, changes in horticultural practices, and development of new technologies such as HTP to accelerate the breeding process should all be explored as options to reduce water use in processing tomato production.

Conclusions and Hypothetical Future Directions for Research

Despite the overall lack of readily useable data obtained from the HTP robot image and point cloud data, the HTP robot used in this study does show initial promise for use in plant breeding. Only a limited amount of RPN data were obtained from the larger image data set collected by the robot. This limited RPN data set did exhibit a strong correlation between RPN at 114 days post-transplant and manually collected RYLD data. Yield is one of the most important traits for tomato growers and a target for breeders (Hartz et al., 2008; Barrios-Masias and Jackson, 2014). The strong correlations between RPN at 114 days post-transplant and RYLD data are a preliminary indication that the HTP robot was able to accurately reflect the manually collected phenotype data.

The effectiveness of the HTP robot at collecting accurate phenotype data for the other traits (Table 1) was not able to be determined at this time. Suitable proxy trait data may be obtained in the future as these data are fully processed. Evaluation of important traits for processing tomato including total yield and maturity (DAPG, DAPR) would be the next step needed in order to compare the HTP robot collected image/sensor data to the manually collected phenotype data. Future iterations of the HTP robot will need to address technical

limitations that prevented data collection in less than optimal field conditions in this experiment.

Supplemental Table Legends:

Supplemental Table 1:

List of genotypes used in each of two phases of the experiment, training and validation. Table A lists the training set germplasm which was composed of introgression lines derived from *S. habrochaites* and inbred cultivars. Table B lists the validation set germplasm composed of introgression lines, inbred cultivars, commercial F1 hybrids, and experimental F1 hybrids.

Supplemental Table 2:

Summary of F test values for ANOVAs performed on trait data for the training set germplasm. An - indicates not included in the model, and a blank space in the field location(s) or water treatment column indicates that data for locations or water treatments were combined.

Supplemental Table 3:

Summary of F test values for ANOVAs performed on trait data for the validation set germplasm. An - indicates not included in the model, and a blank space in the field location(s) or water treatment column indicates that data for locations or water treatments were combined.

Supplemental Table 4:

Trait means and Tukey mean separation groups for the training set germplasm. Lines were evaluated in the field under two water treatments. Genotypes are listed to the left. Under each trait, trait means for each accession are listed to the left. Tukey groups are listed to the right of each trait. An N/A is used to indicate that Orion was not included for that trait-year combination.

Supplemental Table 5:

Trait means and Tukey mean separation groups for the validation set of genotypes. Lines were evaluated in the field under two water treatments. Genotypes are listed to the left. Under each trait, trait means for each accession are listed to the left. Tukey groups are listed to the right of each trait.

Supplemental Table 6:

Summary of F test values for ANOVAs performed on RPN data for the training set of genotypes.

Acknowledgements:

This research was supported by the US Department of Agriculture National Institute of Food and Agriculture competitive grants. We thank the St. Clair lab members for contributions to field and lab data collection. We also thank our collaborators in Dr. David Slaughter's lab in the UC Davis Biological and Agricultural engineering department. Lastly, we thank the UC Davis Plant Sciences field facility crew for assistance with field experiments.

References:

Araus, J.L. and Cairns, J.E. 2014. Field high-throughput phenotyping: the new crop breeding frontier. *Frontiers in Plant Science.* 19(1): 52-61. https://doi.org/10.1016/j.tplants.2013.09.008

Araus, J.L., Kefauver, S.C., Zaman-Allah, M., Olsen, M.S. and Cairns, J.E. 2018. Translating high-throughput phenotyping into genetic gain. *Trends in Plant Science*. 23(5): 451-466. https://doi.org/10.1016/j.tplants.2018.02.001

Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2015. High-resolution mapping of a major effect QTL from wild tomato *Solanum habrochaites* that influences water relations under root chilling. *Theoretical and Applied Genetics*. 128: 1713-1724. https://doi.org/10.1007/s00122-015-2540-y

Atefi, A., Ge, Y., Pitla, S. and Schnable, J. 2021. Robotic technologies for high-throughput plant phenotyping: contemporary reviews and future perspectives. *Frontiers in Plant Science*. 12: 611940. https://doi.org/10.3389/fpls.2021.611940

Bai, G., Ge, Y.F., Hussain, W., Baenziger, P.S. and Graef, G. 2016. A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. *Computers and Electronics in Agriculture*. 128: 181-192. https://doi.org/10.1016/j.compag.2016.08.021

Bai, Y., Kissoudis, C., Yan, Z., Visser, R.G.F. and Van der Linden, G. 2018. Plant behaviour under combined stress: tomato responses to combined salinity and pathogen stress. *The Plant Journal*. 93: 781-793. https://doi.org/10.1111/tpj.13800

Barrios-Masias, F.H. and Jackson, L.E. 2014. California processing tomatoes: morphological, physiological and phenological traits associated with crop improvement during the last 80 years. *European Journal of Agronomy*. 53: 45-55. https://doi.org/10.1016/j.eja.2013.11.007

Bates, D., Maechler, M., Bolker, B., and Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 67(1): 1-48. https://doi.org/10.48550/arXiv.1406.5823

Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science*. 48(5): 1649-1664. https://doi.org/10.2135/cropsci2008.03.0131

Chang, A.J., Jung, J.H., Yeom, J., Maeda, M.M., Landivar, J.A., Enciso, J.M., Avila, C.A. and Anciso, J.R. 2021. Unmanned aircraft system- (UAS-) based high-throughput phenotyping (HTP) for tomato yield estimation. *Journal of Sensors*. 2021: 8875696. https://doi.org/10.1155/2021/8875606

Chetelat, R.T., and Ji, Y. 2007. Cytogenetics and evolution. In *Genetic Improvement of Solanaceous Crops* (Razdan, M.K. and Mattoo, A.K., eds). Science Publishers. Enfield, NH. pp 77-112.

Chetelat, R.T., Pertuze, R.A., Faundez, L., Graham, E.B. and Jones, C.M. 2009. Distribution, ecology, and reproductive biology of wild tomatoes and related nightshades from the Atacama Desert region of northern Chile. *Euphytica*, 167: 77-93. https://doi.org/10.1007/s10681-008-9863-6

Cobb, J.N., DeClerck, G. Greenburg, A., Clark, R. and McCouch, S. 2013. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phentoype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*. 126: 867-887. https://doi.org/10.1007/s00122-013-2066-0

Corrado, G., Piffanelii, P., Caramente, M., Coppola, M. and Rao, R. 2013. SNP genotyping reveals genetic diversity between cultivated landraces and contemporary varieties of tomato. *BMC Genomics*. 14: 835.

Dariva, F.D., Copati, M.G.F., Pessoa, H.P., Alves, F.M., de Oliveira Dias, F., de Toledo Picoli, E.A., da Cunha, F.F. and Nick, C. 2020. Evaluation of anatomical and physiological traits of *Solanum pennellii* Cor. associated with plant yield in tomato plants under water-limited conditions. *Scientific reports*, *10*(1): 1-13. https://doi.org/10.1038/s41598-020-73004-4

Deery, D.M., Rebetzke, G.J., Jimenez-Berni, J.A., James, R.A., Condon, A.G., Bovill, W.D., Hutchinson, P., Scarrow, J., Davy, R. and Furbank, R.T. Methodology for High-Throughput Field Phenotyping of Canopy Temperature Using Airborne Termography. *Frontiers in Plant Science*. 7: 1808. https://doi.org/10.3389/fpls.2016.01808

Deitch, M.J., Sapundjieff, M.J. and Feirer, S.T. 2017. Characterizing precipitation variability and trends in world's Medditerranean-climate areas. *Water*. 9(4): 1-21.

Diffenbaugh, N.S., Swain, D.L. and Touma D. 2015. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*. 112(13): 3931-3936. https://doi.org/10.1073/pnas.1422385112

Fahlgren, N. Gehan, M.A. and Baxter, I. 2015. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current Opinion in Plant Biology*. 24: 93-99. https://doi.org/10.1016/j.pbi.2015.02.006

Fehr, W.R. 1987. *Principles of Cultivar Development, vol. 1: Theory and Technique*. Iowa State University Press. Ames, Iowa. Pages 247-249

Fonteijn, H., Afonso, M., Lensink, D., Mooij, M., Faber, N., Vroegop, G.P. and Wehrens, R. 2021. Automatic phenotyping of tomatoes in production greenhouses using robotics and computer vision: From theory to practice. *Agronomy-Basel*. 11(8): 1599. https://doi.org/10.3390/agronomy11081599

Food and Agriculture Organization of the United Nations. 2021. *FAOSTAT Crops*. FAOSTAT, Rome, Italy. http://www.fao.org/faostat/en/#data/QC (accessed 5 May. 2022).

Furbank, R.T. and Tester, M. 2011. Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*. 16(12): 1360-1385. https://doi.org/10.1016/j.tplants.2011.09.005

Goodstal, F.J., Kohler G.R., Randall L.B., Bloom A.J. and St. Clair D.A. 2005. A major QTL introgressed from wild *Lycopersicon hirsutum* confers chilling tolerance to cultivated tomato (*Lycopersicon escultentum*). *Theoretical and Applied Genetics*. 111: 898-905. https://doi.org/10.1007/s00122-005-0015-2

Haghshenas, A. and Emam, Y. 2019. Evidence for relative grain yield simulation by red color level of beneath-canopy soil at wheat booting phase: An unexpected observation using image processing. *Computers and Electronic in Agriculture*. 162: 1028-1034. https://doi.org/10.1016/j.compag.2019.05.044

Halperin, O., Gebremedhin, A., Wallach, R. and Moshelion, M. 2016. High-throughput physiological phenotyping and screening system for the characterization of plant-environment interactions. *The Plant Journal*. 89: 839-850. https://doi.org/10.1111/tpj.13425

Hartz, T.K., Miyao, G., Mickler., Lestrange, M., Stoddard., Nuñez, J., and Aegerter B. 2008. Processing Tomato Production in California. *University of California Agriculture and Natural Resources*, doi: 10.3733/ucanr.7228

Ispiryan, R., Grigoriev, I., Castell, W.Z. and Schaffner, A.R. 2013. A segmentation procedure using color features applied to images of *Arabidopsis thaliana*. *Functional Plant Biology*. 40(10): 1065-1075. https://doi.org/10.1071/FP12323

Kar, S., Garin, V., Kholova, J., Vadez, V., Durbha, S.S., Tanaka, R., Iwata, H., Urban, M.O. and Adinarayana, J. 2020. SpaTemHTP: A data analysis pipeline for efficient processing and utilization of temporal high-throughput phenotyping data. *Frontiers in Plant Science*. 11: 552508. https://doi.org/10.3389/flps.2020.552509

Kaur, A., Donis-Gonzalez, I.R. and St. Clair, D.A. 2020. Evaluation of a hand-held spectrophotometer as an in-field phenotyping tool for tomato and pepper fruit quality. *The Plant Phenome Journal*. 3(1): e20008. https://doi.org/10.1002/ppj2.20008

Kubond, B.A. and St. Clair, D.A. 2022. Bin mapping of water stress tolerance-related, fruit quality, and horticultural traits in tomato introgression lines derived from wild *Solanum habrochaites*. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20869

Kulus, D. 2018. Genetic resources and selected conservation methods of tomato. *Journal of Applied Botany and Food Quality.* 91: 135-144. https://doi.org/10.5073/JABFQ/2018.091.019

Lee, J., De Gryze, S., Six, J. 2010. Effect of climate change on field crop production in California's Central Valley. *Climate Change*. 109: 335-353. https://doi.org/10.1007/s10584-011-0305-4

Lenth, R.V. 2016. Least-Squares Means: The R package lsmeans. *Journal of Statistical Software*, 69(1): 1-33. https://doi.org/10.18637/jss.v069.i01

Lounsbery, J.K., Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2016. Quantitative Trait Loci for water-stress tolerance traits localize on chromosome 9 of wild tomato. *Crop Science*, 56: 1514-1525. https://doi.org/10.2135/cropsci2015.07.0432

Mahoney, K., Scott, J.D., Alexander, M., McCray, R., Hughes, M., Swales, D. and Bukovsky, M. 2021. Cool season precipitation projections for California and the Western United States in NA-CORDEX models. *Climate Dynamics*. 56: 3081-3102. https://doi.org/10.1007/s00382-021-05632-z

Miller, J.C. and Tanksley, S.D. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theoretical and Applied Genetics*. 80: 437-448. https://doi.org/10.1007/BF00226743

Mote, P.W., Hamlet, A.F., Clark, M.P. and Lettenmaier, D.P. 2005. Declining mountain snowpack in western North America. *Bulletin of the American Meterological Society*. 86(1):39-49. https://doi.org/10.1175/BAMS-86-1-39

Moyle L.C. and Muir C.D. 2010. Reciprocal insights into adaptation from agricultural and evolutionary studies in tomato. *Evolutionary Applications*. 3(5-6): 409-421. https://doi.org/10.1111/j.1752-4571.2010.00143.x

National Agricultural Statistics Service. 2021. *Statistics by subject*. USDA, Washington DC. http://www.nass.usda.gov/Statistics_by_Subject/?sector=CROPS (accessed 5 May 2022).

Pathak, T., Maksey, M.L., Dahlberg, J.A., Kearns, F., Balie, K.M. and Zaccaria, D. 2018. Climate change trends and impacts on California agriculture: a detailed review. *Agronomy*. 8(25): 1-27. https://doi.org/10.3390/agronomy8030025

R Core Team 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Ray, P., Wi, S., Schwarz, A., Correa, M., He, M.X. and Brown, C. 2020. Vulnerability and risk: climate change and water supply from California's central valley water system. *Climate Change*. 161(1): 177-199. https://doi.org/10.1007/s10584-020-02655-z

Rick C.M. 1973. Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. *Genes, Enzymes, and Populations*. 255-269. https://doi.org/10.1007/978-1-4684-2880-3_17

Rick C.M. 1983. Genetic variability in tomato species. *Plant Mol. Biol. Rep.* 1: 81-87. https://doi.org/10.1007/BF02680303

Roo, C. 2022. *Architectural Phenotype Measurements Utilizing High Throughput Phenotyping in Solanaceae Varieties* (Roo_ucdavis_0029M_20990). [Master's Thesis, University of California, Davis]. Escholarship. https://escholarship.org/uc/item/1470b611

Roo, C., Slaughter, D.C. and Vuong, V.L. 2020, July 12-15. *Volume estimation through high throughput phenotyping* [Conference presentation]. American Society of Agricultural and Biological Engineers Annual International Meeting, Omaha, NE, United States. https://elibrary.asabe.org/azdez.asp?JID=5&AID=51584&CID=virt2020&v=&i=&T=1&refer=7&access=&dabs=Y

Schouten, H.J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y. and Visser, R.G.F. 2019. Breeding has increased the diversity of cultivated tomato in The Netherlands. *Frontiers in Plant Science*. 10: 1606. https://doi.org/10.3389/fpls.2019.01606

Seager, R., Osborn, T.J., Kushnir, Y., Simpson, I.R., Nakamura, J. and Liu, H.B. 2019. Climate variability and change of Mediterranean-type climates. *Journal of Climate*. 32(10): 2887-2915. https://doi.org/10.1175/JCLI-D-18-0472.1

Song, P., Wang, J.L., Guo, X.Y., Yang, W.N. and Zhao, C.J. 2021. High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. *Crop Journal*. 9(3): 633-645. https://doi.org/10.1016/j.cj.2021.03.015

Spooner, D.M., Peralta, E. and Knapp, S. 2005. Comparison of AFLPs with other markers for phylogenetic inference in wild tomatoes (*Solanum L.* section *Lycopersicon* (Mill.) Wettst.). *Taxon*. 54:43-61.47. https://doi.org/10.2307/25065301

Tamburino, R., Sannino, L., Cafasso, D., Cantarella, C., Orru, L., Cardi, T., Cozzolino, S., D'Agostino, N. and Scotti, N. 2020. Cultivated tomato (*Solanum lycopersicum* L.) suffered a severe cytoplasmic bottleneck during domestication: implications from chloroplast genomes. *Plants-Basel*. 9(11): 1443. https://doi.org/10.3390/plants9111443

Tester, M. and Langridge, P. 2010. Breeding technologies to increase crop production in a changing world. *Science*. 327(5967): 818-822. https://doi.org/10.1126/science.1183700

Vuong, V.L., Slaughter, D.C., Roo, C., St. Clair, D.A., Kubond, B.A. and Bosland, P. 2020, April 27. *In-field high-throughput phenotyping approach using a multi-view and multi-sensor ground-*

based vehicle [Conference Presentation]. SPIE 11414, Virtual. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11414/114140H/In-field-high-throughput-phenotyping-approach-using-a-multi-view/10.1117/12.2560531.short?SSO=1

Walsh, B. 2009. Quantitative genetics, version 3.0: where have we gone since 1987 and where are we headed? *Genetica*. 136(2): 213-223. https://doi.org/10.1007/s10709-008-9324-0

Wu, X.D., Zhu, X.Q., Wu, G.Q. and Ding, W. 2014. Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*. 26(1): 97-107. https://doi.org/10.1109/TKDE.2013.109

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Zi., Li, H. and Yang, H. 2017. Unmanned aerial vehicle remote sensing for field-based crop phenotyping: current status and perspectives. *Frontiers in Plant Science*. 8: 1111. https://doi.org/10.3389/fpls.2017.01111

Zhang, Q.C., Yang, L.T., Chen, Z.K. and Li, P. 2018. A survey on deep learning for big data. *Information Fusion*. 42: 146-157. https://doi.org/10.1016/j.inffus.2017.10.006

Conclusions and Future Directions:

Introgression lines derived from wild *Solanum habrochaites* were found to exhibit a wealth of trait diversity that may prove useful for the improvement of cultivated tomato. A total of 268 TGRAs were found among 22 of 24 ILs included in the first chapter (Kubond and St. Clair 2022). TGRAs with positive horticultural effects were identified in this study for 7 of 15 traits, Sunburn, Canopy, DAPG, DAPR, FW, Brix, and pH. For these traits, *S. habrochaites* is a possible donor source for the improvement of cultivated tomato.

The ILs selected for use in a combining ability study showed limited potential for use in a hybrid breeding program. Of the five IL parents selected, only LA3933 containing an introgression from *S. habrochaites* on chromosome 4 may be suitable for use in a hybrid breeding program. The other four IL parents exhibited horticulturally undesirable GCA estimates for yield, making them undesirable due to the importance of yield as a trait in tomato (Hartz et al., 2008; Barrios-Masias and Jackson, 2014). LA3933, in contrast to the other IL parents, has predominantly horticulturally positive GCA estimates for TYLD, RYLD, HI, and DAPG. These generally desirable GCA estimates for these key traits suggests that LA3933 is the best candidate for use in a hybrid breeding program. The other ILs evaluated may be useful for the improvement of processing tomato if evaluated for combining ability in other genetic backgrounds besides E6203.

The tractor-based HTP robot showed preliminary promise in accurately phenotyping processing tomato for ripe yield. Despite the small dataset that was available to analyze, a

single main plot at one location in one year of the experiment, a strong Spearman rank correlation of r = 0.76 was found between the manually collected RYLD phenotype data and the HTP robot collected RPN data. Another study using a UAV system also identified strong correlations for yield in tomato (Chang et al., 2021). The strong correlation identified in the limited dataset shows preliminary promise for the use of HTP technologies to accurate collect plant phenotype data.

The reduced water treatment applied in each experiment was not severe enough to induce significant differences in trait genotype means between the two water treatments for the recurrent parent (E6203) control in any of the experiments. The experiments in this dissertation (in contrast to previous studies in the St. Clair lab) used a milder deficit water treatment of 40% Etc applied once fruit set was achieved. This treatment was a relatively mild stress compared to the season long severe deficit 33% Etc treatment used previously in studies by the St. Clair lab (Lounsbery et al., 2016; Groh et al., 2022).

A future experiment that adopts a more severe deficit treatment similar to that used in Lounsbery et al. (2016) or Groh et al. (2022), but otherwise has the same experimental structure as chapter 1 is a possible next step for this research. A more severe deficit water treatment may result in significant mean differences between E6203 and introgression lines for various traits, particularly yield between the two water treatments. A more severe treatment may lead to the identification of TGRAs for one of the water-stress tolerance related traits and/or yield that could be beneficial towards the improvement of cultivated tomato.

Additionally, since the experiments included in the first chapter were conducted, a new IL library for *S. sitiens* was developed and is available (Chetelat et al., 2019). *S. sitiens* is a tomato wild relative native to the Atacama Desert of Chile, and is tolerant to the extreme dry conditions of the Atacama (Rick 1973; Chetelat et al., 2019). A similar experiment to the one performed in chapter 1 using ILs from the *S. sitiens* IL library would likely lead to the identification of TGRAs for water-stress tolerance-related traits that could be used to improve cultivated tomato tolerance to water limited conditions. This experiment would also utilize a more restrictive water-stress treatment similar to that used in Lounsbery et al. (2016) or Groh et al. (2022) to help reveal statistically significant differences between water treatments. If the *S. sitiens* IL library had been available prior to the start of my PhD, it likely would have been a better choice for the first chapter if my main aim were to improve only water-stress tolerance-related traits.

To validate the overall effectiveness of the HTP robot, the remainder of the data collected by the HTP robot needs to be processed and then compared to the manually collected phenotype data. The number of plants included in the experiments for chapter 3 was significantly higher than in previous HTP studies on tomato, approximately 20,000 plants vs 100-400 plants (Halperin et al., 2016; Chang et al., 2021; Fonteijn et al., 2021). Processing the image and time-of-flight data from the HTP robot into a form that can be compared to the manual phenotype data would be the next step to determining the effectiveness of the HTP robot at collecting accurate phenotype data. Once suitable data from the HTP robot is obtained

for comparison it can be compared to the manually collected phenotype data for the remaining 10 traits as well as RYLD.

The limited dataset for comparison of RYLD to RPN presented in chapter 3 is not sufficient to make overall conclusions about the effectiveness of the HTP robot. HTP robot data was only available for a single main plot in a single location in a single year. The phenotype for a particular trait is a result of the combination of genetics, environment, and the interaction between the two (Fehr 1987; Chang et al., 2021). The single environment in which RPN data was obtained limits the scope of the conclusions that can be made about the efficacy of the HTP robot for collecting accurate ripe yield data. Plant breeders evaluate material over a range of target environments when making evaluations of material (Fehr 1987; Bernardo 2008; Chang et al., 2021). Further trialing of the HTP robot at multiple locations would be needed to determine its overall effectiveness at collecting representative phenotype data from tomato. Four locations of training data were collected by the HTP robot across two years in addition to two locations of validation data in a single year. The current iteration of the HTP robot is a prototype that is not designed for use in sub-optimal field conditions that can be present in plant breeder's fields. Further development of the HTP robot to allow for entry into fields with sub-optimal conditions would be needed prior to use in tomato breeding. Processing the HTP robot data for the entire dataset and comparing it to the manually collected phenotype data would be needed to further assess the ability of the HTP robot to collect accurate phenotype data in more than a single environment.

Conclusion and Future Directions References:

Barrios-Masias, F.H. and Jackson, L.E. 2014. California processing tomatoes: morphological, physiological and phenological traits associated with crop improvement during the last 80 years. *European Journal of Agronomy*. 53: 45-55. https://doi.org/10.1016/j.eja.2013.11.007

Bernardo, R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Science*. 48(5): 1649-1664. https://doi.org/10.2135/cropsci2008.03.0131

Chang, A.J., Jung, J.H., Yeom, J., Maeda, M.M., Landivar, J.A., Enciso, J.M., Avila, C.A. and Anciso, J.R. 2021. Unmanned aircraft system- (UAS-) based high-throughput phenotyping (HTP) for tomato yield estimation. *Journal of Sensors*. 2021: 8875696. https://doi.org/10.1155/2021/8875606

Chetelat, R.T., Qin, X., Tan, M., Burkart-Waco, D., Moritama, Y., Huo, X., Wills, T. and Pertuzé R. 2019. Introgression lines of *Solanum sitiens*, a wild nightshade of the Atacama Desert, in the genome of cultivated tomato. *The Plant Journal*. 100(4): 836-850. https://doi.org/10.1111/tpj.14460

Fehr, W.R. 1987. *Principles of Cultivar Development, vol. 1: Theory and Technique*. Iowa State University Press. Ames, Iowa. Pages 428-434

Fonteijn, H., Afonso, M., Lensink, D., Mooij, M., Faber, N., Vroegop, G.P. and Wehrens, R. 2021. Automatic phenotyping of tomatoes in production greenhouses using robotics and computer vision: From theory to practice. *Agronomy-Basel*. 11(8): 1599. https://doi.org/10.3390/agronomy11081599

Groh, A.M., Kubond, B.A., and St. Clair, D.A. 2022. Fine mapping of QTL for water use efficiency-related traits on chromosome 9 of *Solanum habrochaites* in the field. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20828

Halperin, O., Gebremedhin, A., Wallach, R. and Moshelion, M. 2016. High-throughput physiological phenotyping and screening system for the characterization of plant-environment interactions. *The Plant Journal*. 89: 839-850. https://doi.org/10.1111/tpj.13425

Hartz, T.K., Miyao, G., Mickler., Lestrange, M., Stoddard., Nuñez, J., and Aegerter B. 2008. Processing Tomato Production in California. *University of California Agriculture and Natural Resources*, doi: 10.3733/ucanr.7228

Kubond, B.A. and St. Clair, D.A. 2022. Bin mapping of water stress tolerance-related, fruit quality, and horticultural traits in tomato introgression lines derived from wild *Solanum habrochaites*. *Crop Science*. Advance Online Publication. https://doi.org/10.1002/csc2.20869

Lounsbery, J.K., Arms, E.M., Bloom, A.J. and St. Clair, D.A. 2016. Quantitative Trait Loci for water-stress tolerance traits localize on chromosome 9 of wild tomato. *Crop Science*, 56: 1514-1525. https://doi.org/10.2135/cropsci2015.07.0432

Rick C.M. 1973. Potential Genetic Resources in Tomato Species: Clues from Observations in Native Habitats. *Genes, Enzymes, and Populations*. 255-269. https://doi.org/10.1007/978-1-4684-2880-3_17

Appendix:

Chapter 1 Supplemental Tables

Supplemental Table 1: Genotype x Environment interactions detected in traits evaluated in a set of 24 ILs derived from *S. habrochaites* grown in the field under two water treatments.

Supplemental Table 2: Trait means and Tukey mean separation groups for 24 ILs derived from *S. habrochaites.*

Supplemental Table 3: Spearman rank correlations for each pairwise combination of categorical traits (Canopy, Habit, and Sunburn) in the ILs plus E6203.

Supplemental Table 4: Detected TGRAs in 24 ILs derived from *S. habrochaites* and evaluated in the field under two water treatments in three years.

Chapter 2 Supplemental Tables

Supplemental Table 1: Summary of F test values for ANOVAs performed on trait data for the set of 15 tomato F1 hybrids, 5 IL parents, 3 tomato inbred line parents, and recurrent parent control E6203 evaluated in the field under two water treatments (full and reduced).

Supplemental Table 2: Trait means and Tukey mean separation groups for 15 tomato F1 hybrids, 5 IL parents, 3 tomato inbred line parents, and recurrent parent control E6203.

Supplemental Table 3: GCA estimates for 5 IL female parents and 3 tomato inbred line male parents.

Supplemental Table 4: SCA estimates for 15 tomato F1 hybrids.

Supplemental Table 5: Heritability estimates, broad and narrow, were calculated for each trait.

Chapter 3 Supplemental Tables

Supplemental Table 1: List of genotypes used in each of two phases of the experiment, training and validation. Table A lists the training set germplasm which was composed of introgression lines derived from *S. habrochaites* and inbred cultivars. Table B lists the validation set germplasm composed of introgression lines, inbred cultivars, commercial F1 hybrids, and experimental F1 hybrids.

Supplemental Table 2: Summary of F test values for ANOVAs performed on trait data for the training set germplasm. An - indicates not included in the model, and a blank space in the field location(s) or water treatment column indicates that data for locations or water treatments were combined.

Supplemental Table 3: Summary of F test values for ANOVAs performed on trait data for the validation set germplasm. An - indicates not included in the model, and a blank space in the field location(s) or water treatment column indicates that data for locations or water treatments were combined.

Supplemental Table 4: Trait means and Tukey mean separation groups for the training set germplasm. Lines were evaluated in the field under two water treatments. Genotypes are listed to the left. Under each trait, trait means for each accession are listed to the left. Tukey groups are listed to the right of each trait. An N/A is used to indicate that Orion was not included for that trait-year combination.

Supplemental Table 5: Trait means and Tukey mean separation groups for the validation set germplasm. Lines were evaluated in the field under two water treatments. Genotypes are listed to the left. Under each trait, trait means for each accession are listed to the left. Tukey groups are listed to the right of each trait.

Supplemental Table 6: Summary of F test values for ANOVAs performed on RPN data for the training germplasm.