UNIVERSIDADE FEDERAL DE VIÇOSA

HERIKA PAULA PESSOA

WILD TOMATO SPECIE *Solanum pennellii* AS A GENETIC RESOURCE TO DISCOVER GENES OF INTEREST TO TOMATO BREEDING

VIÇOSA – MINAS GERAIS
2022

HERIKA PAULA PESSOA

WILD TOMATO SPECIE *Solanum pennellii* AS A GENETIC RESOURCE TO DISCOVER GENES OF INTEREST TO TOMATO BREEDING

Thesis submitted to the Plant Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Carlos Nick Gomes

Co-advisers: Fernando França da Cunha
Pedro Crescêncio Souza Carneiro

VIÇOSA – MINAS GERAIS
2022

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T
P475w
2022

Pessoa, Herika Paula, 1993-
Wild tomato specie *Solanum pennellii* as a genetic resource to discover genes of interest to tomato breeding / Herika Paula Pessoa. – Viçosa, MG, 2022.
1 tese eletrônica (81 f.): il. (algumas color.).

Texto em inglês.
Orientador: Carlos Nick Gomes.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Agronomia, 2022.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2022.378
Modo de acesso: World Wide Web.

1. Tomate - Melhoramento genético. 2. Introgressão. 3. Tomate - Resistência à seca. 4. Sementes - Qualidade. 5. Recursos genéticos vegetais. 6. *Solanum pennellii*. 7. *Solanum lycopersicum*. I. Gomes, Carlos Nick, 1979-. II. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Fitotecnia. III. Título.

CDD 22. ed. 631.52

Bibliotecário(a) responsável: Euzebio Luiz Pinto CRB 6/3317

HERIKA PAULA PESSOA

WILD TOMATO SPECIE *Solanum pennellii* AS A GENETIC RESOURCE TO DISCOVER GENES OF INTEREST TO TOMATO BREEDING

Thesis submitted to the Plant Science Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: March 18, 2022.

Assent:

Herika Paula Pessoa
Author

Carlos Nick Gomes
Adviser

To God and to my parents.

ACKNOWLEDGEMENTS

To God.

To my parents for the unconditional love and support.

To all my family, especially my spouse Ramon for the love, support, encouragement, comfort in the bad days, and celebration in the good daysTo the Federal University of Viçosa, for the opportunity to complete the postgraduate course.

To all my Friends for making my journey more joyful and happy.

To my adviser and now friend Carlos Nick Gomes for wisely guide my graduate studies.

To all the the professors in the Plant Science department, and to all the teachers and professors that crossed my path since my alphabetization until grad school.

To professor Dr. Kevin Crosby for giving me the opportunity to do part of my research at Texas A&M University, and to my colleague and friend Dr. Sixto Marques whose help made this time abroad much easier.

To all the workers of the Research and Extension Farm Unit Horta Velha at the Universidade Federal de Viçosa for the help in installing and conducting the field experiments.

To all the colleagues and friends from NEO and NEPFIT for the incredible support, and to have helped me to develop skills beyond the grad school requirements.

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), for granting the scholarship.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

“knowledge is power”.
(Francis Bacon)

ABSTRACT

PESSOA, Herika Paula, D.Sc., Universidade Federal de Viçosa, March, 2022. **Wild tomato specie S*olanum pennellii* as a genetic resource to discover genes of interest to tomato breeding**. Adviser: Carlos Nick Gomes. Co-advisers: Fernando França da Cunha and Pedro Crescêncio de Souza Carneiro.

The *Solanum pennellii* introgression line (IL) population is a valuable source of gene reservoirs for many desirable traits for tomato breeding. Our main goal was to investigate whether *S. pennellii* introgression fragments carry advantageous genes to increase germination, seedling vigor, and drought tolerance during the vegetative/reproductive stage. If so, we aimed to select promising ILs possessing these traits to be included in tomato breeding programs and investigate their *S. pennellii* fragments to uncover candidate genes associated with seed germination, seedling vigor, and drought tolerance during the vegetative/reproductive stage. A multi-trait index based on factor analysis and genotype–ideotype distance (FAI-BLUP index) was used to rank the ILs according to seed quality and drought tolerance during the vegetative/reproductive stage. Then, introgressed segments of interest were identified for candidate gene identification using the map of the tomato IL population, as provided by the SOL Genomics Network. The FAI-BLUP index ranked the ILs above their parent cultivar M-82 for seed germination, seedling vigor, and drought tolerance during the vegetative/reproductive stage, confirming that fragments of the *S. pennellii* genome can donate genes of interest for these traits. IL 6-1, IL 4-1-1, IL 4-4, IL 11-4-1, IL 1-1-2, and IL 10-1 were closest to the ideotype of seed germination performance and vigor. IL 1-4-18, IL 7-4-1, IL 7-1, IL 7-5-5, and IL 1-2 were ranked closest to the drought-tolerant ideotype during the vegetative/reproductive stage. The genes At1g75350, At2g26350, At2g28490, At3g10920, At5g13200, At5g41480, and At4g21800 are candidate genes for seed germination; At5g53000, At1g46480, At1g30755, At5g47390, At2g29630, and At4g03400 are candidate genes for seedling vigor; and At3g03790, At3g06050, At3g12630, At3g15530, At3g23590, At3g49480, At3g61140, At3g62130, At3g62010, and At3g63520 are candidate genes for drought tolerance in the vegetative/reproductive stage. Because they are already associated with genetic markers, they can be transferred to elite tomato cultivars through marker-assisted technology after validation.

Keywords: *Solanum lycopersicum*. Seed quality. Drought. Water stress. Introgression line. FAI-BLUP.

RESUMO

PESSOA, Herika Paula, D.Sc., Universidade Federal de Viçosa, março de 2022. **Parente silvestre S*olanum pennellii* como recurso genético para descoberta de genes de interesse para o melhoramento de tomateiro**. Orientador: Carlos Nick Gomes. Coorientadores: Fernando França da Cunha e Pedro Crescêncio de Souza Carneiro.

A população de linhagens de introgressão (IL) de *Solanum pennellii* é uma valiosa fonte de reservatório de genes de características desejáveis para o melhoramento do tomateiro. Nosso principal objetivo foi investigar se os fragmentos de *S. pennellii* introgredidos carregam genes que proporcionam aumento da germinação e vigor de sementes e tolerância ao déficit hídrico durante o estádio vegetativo e reprodutivo. Em caso afirmativo, selecionar ILs promissoras que possuam essas características para serem incluídas em nosso programa de melhoramento de tomate e investigá-las para descobrir genes candidatos associados à germinação e vigor de sementes e tolerância ao déficit hídrico durante o estádio vegetativo e reprodutivo. Usamos uma abordagem multicaracterística (índice de seleção FAI-BLUP) para ranquear ILs quanto a germinação e vigor de sementes e tolerância ao déficit hídrico durante o estágio vegetativo/reprodutivo. Em seguida, os fragmentos de interesse foram investigados para identificação de genes candidatos usando o mapa da população de linhagens de introgressão disponíveis na plataforma SOL Genomics Network. O índice FAI-BLUP ranqueou ILs superiores a cultivar M-82 tanto para germinação e vigor como para tolerância ao déficit hídrico confirmando que os fragmentos do genoma de *S. pennellii* podem doar genes de interesse para essas características. IL 6-1, 4-1-1, 4-4, IL 11-4-1, IL 1-1-2 e IL 10-1 foram ranqueadas como os genótipos como mais próximos ao ideótipo de germinação e vigor de sementes. IL 1-4-18, IL 7-4-1, IL 7-1, IL 7-5-5 e IL 1-2 foram ranqueadas como genótipos mais próximos ao ideótipo de tolerância ao déficit hídrico no estágio vegetativo/reprodutivo. Os genes At1g75350, At2g26350, At2g28490, At3g10920, At5g13200, At5g41480 e At4g21800 são genes candidatos para germinação de sementes. Os genes At5g53000, At1g46480, At1g30755, At5g47390, At2g29630 e At4g03400 são genes candidatos para o vigor da semente. Os genes At1g55870, At1g55840, At3g06050, At3g12630, At3g23590, At3g26060, At3g43230, At3g62130, At3g63520 e At4g38810 são genes candidatos para tolerância ao déficit hídrico durante o estágio vegetativo/reprodutivo. Após sua validação, por já estarem associados a marcadores genéticos, esses genes poderão ser transferidos para cultivares elite de tomateiro por meio de tecnologia assistida por marcadores.

Palavras-chave: *Solanum lycopersicum*. Qualidade de sementes. Seca. Déficit hídrico. Linhagens de introgressão. FAI-BLUP.

LIST OF ILLUSTRATIONS

CHAPTER I

CHAPTER II

LIST OF TABLES

CHAPTER I

CHAPTER II

LIST OF SUPPLEMENTARY MATERIAL

CHAPTER I

CHAPTER II

SUMMARY

GENERAL INTRODUCTION

The tomato originated in the South American Andes in a region ranging from northern Chile, through Bolivia and Peru, to Ecuador and Colombia (Bai & Lindhout, 2007). Wild tomatoes are found in a wide range of habitats in western South America, from sea level to elevations above 3600 m (Peralta et al., 2008). Wild species possess several morphophysiological adaptations to survive these diverse environmental conditions (Moyle & Muir, 2010). Tomato domestication is linked to two species—the wild *Solanum pimpinellifolium* L. and the wild and semi-domesticated *Solanum lycopersicum* var. *cerasiforme*—and seems to have followed a two-step process; the first domestication occurred in South America and the second in Mesoamerica (Blanca et al., 2021).

It is believed that Spanish explorers discovered tomatoes in Mexico and Central America, introduced them to Spain, and distributed them throughout their colonies, from where they spread globally (Blanca et al., 2015). Currently, the domesticated tomato, *Solanum lycopersicum* L., is an extremely popular vegetable crop produced and consumed worldwide and is one of the world's most economically important vegetables. Over the past 30 years (1991–2020), the global production area of tomatoes expanded from 2.86–5.03 million hectares and the quantity more than doubled from 76.09–180 million tons (FAO, 2021).

Similar to many autogamous crops, selection and breeding processes have led to a dramatic reduction in genetic variability. Despite the high variability found in morphological analyses, molecular marker analyses have revealed limited genetic variation (approximately 5 %) within cultivated tomatoes (Clint Nesbitt & Tanksley, 2002; Ebert & Schafleitner, 2015). Consequently, genetic variability for traits of interest, such as yield, fruit quality, flavor, nutritional content, seed germination and vigor, tolerance to abiotic stresses such as heat, drought, and salinity, and resistance to biotic stresses, is limited in cultivated tomatoes (Bretó et al., 1993). The natural variation found in wild relatives has been exploited to reintroduce lost genetic variation into several cultivated species (Geshnizjani et al., 2020); however, the introgression of traits from distant wild relatives to cultivated tomatoes has proven difficult because of linkage drag and crossability barriers. For example, *Solanum chilense* can act as a pollen parent for *S. lycopersicum*, but viable seeds are rare and the reciprocal cross is not possible (Raduski & Igić, 2021). *Solanum habrochaites* can act as a pollen parent in crosses with cultivated tomatoes, but the reciprocal cross does not set fruit, and *Solanum peruvianum*

also displays severe crossing barriers in hybridization attempts with cultivated tomatoes (Tyagi et al., 2013).

Solanum pennellii, however, is a crossable wild tomato relative that has caught the attention of breeders as a potential source of several useful traits, including drought tolerance (Bolger et al., 2014). A major technical breakthrough was the development of a collection of *S. pennellii* introgression lines (ILs) for tomatoes (Eshed & Zamir, 1995). The collection consisted of 76 lines containing defined genomic segments of *S. pennellii* (LA0716) replacing homologous regions in a tomato (*S. lycopersicum* cv. M-82) background. The *S. pennellii* segments overlap between the lines, covering the whole genome of the species in the 76 lines.

Since their creation, these ILs have been extensively tested for a large number of traits, leading to the mapping of quantitative trait loci (QTL) associated with yield (J. P. Fernandez-Moreno et al., 2017; Fridman et al., 2004), fruit weight, soluble solids (Baxter et al., 2005; Overy et al., 2005), flavor, fruit firmness, color (Chapman et al., 2012; Dariva et al., 2021; Yang et al., 2016), primary metabolites (Schauer et al., 2006), morphology (Fanourakis et al., 2015), enhanced antioxidant content (Aliberti et al., 2020; J. P. Fernandez-Moreno et al., 2017; Rousseaux et al., 2005), biotic stress resistance (Fernandez-Moreno et al., 2017; Smeda et al., 2018), and drought tolerance. The QTL information obtained from the *S. pennellii* IL population is summarized in the Real-Time QTL Database (http://zamir.sgn.cornell.edu/Qtl/Html/home.htm).

ILs are useful for QTL mapping and gene identification because they carry a single introgressed region in the genome of a cultivated species. This largely avoids the common epistatic biases inherent in mapping populations derived from different species (Lippman et al., 2007). The phenotypic variation in the ILs can be linked with their individual introgression segments, allowing the pyramiding of various quantitative traits from different ILs into new breeding lines to maximize yield, improve tolerance to abiotic stresses, or enhance any other desirable traits (Foolad & Panthee, 2012).

Tomato breeders are constantly working to develop new tomato varieties with high yield potential, resistance to biotic stresses, and desirable fruit quality traits such as flavor and texture (Schouten et al., 2019). However, to meet farmers' needs, two other traits have been called to the attention of tomato breeding programs: seed quality, represented by seed germination and seedling vigor, and drought tolerance.

The ability of seeds to germinate quickly and uniformly and seedling vigor are essential seed quality attributes (Shrestha et al., 2016). These parameters may severely affect seedling

establishment and plant growth; hence, improving these parameters will improve crop production (Finch-Savage & Bassel, 2016). According to Geshnizjani et al. (2020), improving seed quality is essential for sustainable and profitable crop production; therefore, it should be a target of tomato breeding programs. Modern tomato cultivars are sensitive to non-optimal germination conditions, which reduces seedling vigor and limits production to only optimal environments (Foolad et al., 2003); thus, revisiting wild relatives to discover genetic variability in this trait is a suitable strategy.

Furthermore, tomatoes are sensitive to drought stress throughout their developmental stages from seed germination to harvest (Foolad et al., 2007). Although tomatoes are mainly produced under irrigated management systems, even short drought periods can lead to qualitative and quantitative losses (George et al., 2013; Rigano et al., 2016). Hence, breeding drought-tolerant materials is considered a low-cost strategy to ensure crop yield when water availability is limited or unpredictable (Bernier et al., 2008). According to Foolad (2007), genetic variation in drought tolerance is limited within *S. lycopersicum*. Therefore, wild tomato species, such as *S. peruvianum, S. chilense* (Tapia et al., 2016), and *S. pennellii* (Bolger et al., 2014), are considered important sources of drought-tolerant genes.

Because *S. pennellii* is a well-known source of genetic variation for several traits (Bolger et al., 2014; Calafiore et al., 2019; Fernandez-Moreno et al., 2017; Frary et al., 2010; Lawson et al., 1997; Rousseaux et al., 2005; Steinhauser et al., 2011; Yang et al., 2016), we believe that the *S. pennellii* IL population is a valuable genetic resource that should be tested for germination and seedling vigor, as well as drought tolerance during the vegetative and reproductive stages. In addition, as breeding initiatives involve first finding genetic materials carrying genes of interest, followed by transferring these genes into elite germplasms, our main goals are as follows:

1- To investigate whether the *S. pennellii* introgression fragments donate advantageous genes that increase seed quality and drought tolerance.
2- If so, to select promising ILs with these traits to be included in tomato breeding programs and investigate their *S. pennellii* fragments to uncover candidate genes associated with seed quality and drought tolerance.

REFERENCES

Aliberti, A., Olivieri, F., Graci, S., Rigano, M. M., Barone, A., & Ruggieri, V. (2020). Genomic Dissection of a Wild Region in a Superior Solanum pennellii Introgression Sub-Line with High Ascorbic Acid Accumulation in Tomato Fruit. *Genes 2020, Vol. 11, Page 847*, *11*(8), 847. https://doi.org/10.3390/GENES11080847

Bai, Y., & Lindhout, P. (2007). Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future? *Annals of Botany*, *100*(5), 1085. https://doi.org/10.1093/AOB/MCM150

Baxter, C. J., Sabar, M., Quick, W. P., & Sweetlove, L. J. (2005). Comparison of changes in fruit gene expression in tomato introgression lines provides evidence of genome-wide transcriptional changes and reveals links to mapped QTLs and described traits. *Journal of Experimental Botany*, *56*(416), 1591–1604. https://doi.org/10.1093/JXB/ERI154

Bernier, J., Atlin, G. N., Serraj, R., Kumar, A., & Spaner, D. (2008). Breeding upland rice for drought resistance. Journal of the Science of Food and Agriculture, 88(6), 927-939.

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M. J., Francis, D., Causse, M., van der Knaap, E., & Cañizares, J. (2015). Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. *BMC Genomics*, *16*(1), 257. https://doi.org/10.1186/s12864-015-1444-1

Blanca, J., Sanchez-Matarredona, D., Ziarsolo, P., Montero-Pau, J., Knaap, E. van der, Díez, M. J., & Cañizares, J. (2021). Haplotype analyses reveal novel insights into tomato history and domestication including long-distance migrations and latitudinal adaptations. *BioRxiv*, 2021.06.18.448912. https://doi.org/10.1101/2021.06.18.448912

Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., Quesneville, H., Alseekh, S., Sørensen, I., Lichtenstein, G., Fich, E. A., Conte, M., Keller, H., Schneeberger, K., Schwacke, R., Ofner, I., Vrebalov, J., Xu, Y., Osorio, S., … Fernie, A. R. (2014). The genome of the stress-tolerant wild tomato species Solanum pennellii. *Nature Genetics*, *46*(9), 1034–1038. https://doi.org/10.1038/ng.3046

Bretó, M. P., Asins, M. J., & Carbonell, E. A. (1993). Genetic variability in Lycopersicon species and their genetic relationships. *Theoretical and Applied Genetics 1993 86:1*, *86*(1), 113–120. https://doi.org/10.1007/BF00223815

Calafiore, R., Aliberti, A., Ruggieri, V., Olivieri, F., Rigano, M. M., & Barone, A. (2019). Phenotypic and molecular selection of a superior solanum pennellii introgression sub-line suitable for improving quality traits of cultivated tomatoes. *Frontiers in Plant Science*, *10*, 190. https://doi.org/10.3389/FPLS.2019.00190/BIBTEX

Chapman, N. H., Bonnet, J., Grivet, L., Lynn, J., Graham, N., Smith, R., Sun, G., Walley, P. G., Poole, M., Causse, M., King, G. J., Baxter, C., & Seymour, G. B. (2012). High-Resolution Mapping of a Fruit Firmness-Related Quantitative Trait Locus in Tomato Reveals Epistatic Interactions Associated with a Complex Combinatorial Locus. *Plant Physiology*, *159*(4), 1644–1657. https://doi.org/10.1104/PP.112.200634

Clint Nesbitt, T., & Tanksley, S. D. (2002). Comparative Sequencing in the Genus Lycopersicon: Implications for the Evolution of Fruit Size in the Domestication of

Cultivated Tomatoes. *Genetics*, *162*(1), 365–379. https://doi.org/10.1093/GENETICS/162.1.365

Dariva, F. D., Pessoa, H. P., Copati, M. G. F., de Almeida, G. Q., de Castro Filho, M. N., Picoli, E. A. de T., da Cunha, F. F., & Nick, C. (2021). Yield and fruit quality attributes of selected tomato introgression lines subjected to long-term deficit irrigation. *Scientia Horticulturae*, *289*, 110426. https://doi.org/10.1016/J.SCIENTA.2021.110426

Ebert, A. W., & Schafleitner, R. (2015). *Utilization of Wild Relatives in the Breeding of Tomato and Other Major Vegetables*.

Eshed, Y., & Zamir, D. (1995). An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield- associated QTL. *Genetics*, *141*(3), 1147–1162.

FAO (2021). Food and Agriculture Organization of the Union. https://doi.org/10.4060/cb2395fr (Consulted on 10/02/2022).

Fanourakis, D., Giday, H., Milla, R., Pieruschka, R., Kjaer, K. H., Bolger, M., Vasilevski, A., Nunes-Nesi, A., Fiorani, F., & Ottosen, C. O. (2015). Pore size regulates operating stomatal conductance, while stomatal densities drive the partitioning of conductance between leaf sides. *Annals of Botany*, *115*(4), 555–565. https://doi.org/10.1093/AOB/MCU247

Fernandez-Moreno, J. P., Levy-Samoha, D., Malitsky, S., Monforte, A. J., Orzaez, D., Aharoni, A., & Granell, A. (2017). Uncovering tomato quantitative trait loci and candidate genes for fruit cuticular lipid composition using the Solanum pennellii introgression line population. *Journal of Experimental Botany*, *68*(11), 2703–2716. https://doi.org/10.1093/JXB/ERX134

Finch-Savage, W. E., & Bassel, G. W. (2016). Seed vigour and crop establishment: Extending performance beyond adaptation. In *Journal of Experimental Botany* (Vol. 67, Issue 3, pp. 567–591). Oxford University Press. https://doi.org/10.1093/jxb/erv490

Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, *2007*. https://doi.org/10.1155/2007/64358

Foolad, M. R., & Panthee, D. R. (2012). Marker-Assisted Selection in Tomato Breeding. *Https://Doi.Org/10.1080/07352689.2011.616057*, *31*(2), 93–123. https://doi.org/10.1080/07352689.2011.616057

Foolad, M. R., Subbiah, P., & Zhang, L. (2007). Common QTL affect the rate of tomato seed germination under different stress and nonstress conditions. *International Journal of Plant Genomics*, *2007*. https://doi.org/10.1155/2007/97386

Foolad, M. R., Zhang, L. P., & Subbiah, P. (2003). Genetics of drought tolerance during seed germination in tomato: inheritance and QTL mapping. *Genome*, *46*(4), 536–545. https://doi.org/10.1139/g03-035

Frary, A., Göl, D., Keleş, D., Ökmen, B., Pınar, H., Şığva, H. Ö., Yemenicioğlu, A., & Doğanlar, S. (2010). Salt tolerance in Solanum pennellii: antioxidant response and related QTL. *BMC Plant Biology*, *10*(1), 58. https://doi.org/10.1186/1471-2229-10-58

Fridman, E., Carrari, F., Liu, Y. S., Fernie, A. R., & Zamir, D. (2004). Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science*, *305*(5691), 1786–1789. https://doi.org/10.1126/science.1101666

George, S., Ahmad Jatoi, S., & Uddin Siddiqui, S. (2013). Genotypic differences against PEG simulated drought stress in tomato. *Pakistan Journal of Botany*, *45*(5), 1551–1556.

Geshnizjani, N., Snoek, B. L., Willems, L. A. J., Rienstra, J. A., Nijveen, H., Hilhorst, H. W. M., & Ligterink, W. (2020). Detection of QTLs for genotype × environment interactions in tomato seeds and seedlings. *Plant Cell and Environment*, *43*(8), 1973–1988. https://doi.org/10.1111/pce.13788

Lawson, D. M., Lunde, C. F., & Mutschler, M. A. (1997). Marker-assisted transfer of acylsugar-mediated pest resistance from the wild tomato, Lycopersicon pennellii, to the cultivated tomato, Lycopersicon esculentum. *Molecular Breeding 1997 3:4*, *3*(4), 307–317. https://doi.org/10.1023/A:1009677412902

Lippman, Z. B., Semel, Y., & Zamir, D. (2007). An integrated view of quantitative trait variation using tomato interspecific introgression lines. In *Current Opinion in Genetics and Development* (Vol. 17, Issue 6, pp. 545–552). Elsevier Current Trends. https://doi.org/10.1016/j.gde.2007.07.007

Moyle, L. C., & 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

Overy, S. A., Walker, H. J., Malone, S., Howard, T. P., Baxter, C. J., Sweetlove, L. J., Hill, S. A., & Quick, W. P. (2005). Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. *Journal of Experimental Botany*, *56*(410), 287–296. https://doi.org/10.1093/JXB/ERI070

Peralta, I. E., Spooner, D. M., & Knapp, S. (2008). Taxonomy of wild tomatoes and their relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae). *Systematic Botany Monographs*, *84*.

Raduski, A. R., & Igić, B. (2021). Biosystematic studies on the status of Solanum chilense. *American Journal of Botany*, *108*(3), 520–537. https://doi.org/10.1002/AJB2.1621

Rigano, M. M., Arena, C., di Matteo, A., Sellitto, S., Frusciante, L., & Barone, A. (2016). Eco-physiological response to water stress of drought-tolerant and drought-sensitive tomato genotypes. *Plant Biosystems*, *150*(4), 682–691. https://doi.org/10.1080/11263504.2014.989286

Rousseaux, M. C., Jones, C. M., Adams, D., Chetelat, R., Bennett, A., & Powell, A. (2005). QTL analysis of fruit antioxidants in tomato using Lycopersicon pennellii introgression lines. *Theoretical and Applied Genetics*, *111*(7), 1396–1408. https://doi.org/10.1007/S00122-005-0071-7/FIGURES/5

Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Bruedigam, C., Kopka, J., Willmitzer, L., Zamir, D., & Fernie, A. R. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology 2006 24:4*, *24*(4), 447–454. https://doi.org/10.1038/nbt1192

Schouten, H. J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y., & 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

Shrestha, S., Knapič, M., Žibrat, U., Deleuran, L.C., Gislum, R., 2016. Single seed near-infrared hyperspectral imaging in determining tomato (Solanum lycopersicum L.) seed quality in association with multivariate data analysis. Sensors Actuators, B Chem. 237, 1027–1034. https://doi.org/10.1016/j.snb.2016.08.170

Smeda, J. R., Schilmiller, A. L., Anderson, T., Ben-Mahmoud, S., Ullman, D. E., Chappell, T. M., Kessler, A., & Mutschler, M. A. (2018). Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Molecular Breeding*, *38*(1), 1–20. https://doi.org/10.1007/S11032-017-0756-Z/FIGURES/2

Steinhauser, M.-C., Steinhauser, D., Gibon, Y., Bolger, M., Arrivault, S., Usadel, B., Zamir, D., Fernie, A. R., & Stitt, M. (2011). Identification of Enzyme Activity Quantitative Trait Loci in a Solanum lycopersicum x Solanum pennellii Introgression Line Population. *Plant Physiology*, *157*(3), 998–1014. https://doi.org/10.1104/pp.111.181594

Tapia, G., Méndez, J., & Inostroza, L. (2016). Different combinations of morpho-physiological traits are responsible for tolerance to drought in wild tomatoes Solanum chilense and Solanum peruvianum. *Plant Biology*, *18*(3), 406–416. https://doi.org/10.1111/PLB.12409

Tyagi, W., Rai, M., & Ngasepam, T. M. (2013). *Wide hybridization and embryo-rescue for crop improvement in Solanum Quantum dots for sensing applications View project Golden rice project View project*. https://doi.org/10.4172/2168-9881.S11-004

Yang, S., Yu, Q., Wang, B., Yang, T., Li, N., Tang, Y., Aisimutuola, P., Wang, Q., Xu, J., & Gao, J. (2016). Identification of QTLs for red fruit firmness using the wild tomato species Solanum pennellii LA716 introgression lines. *Plant Breeding*, *135*(6), 728–734. https://doi.org/10.1111/PBR.12423

CHAPTER I - Uncovering tomato candidate genes for seed germination performance and seedling vigor using the *Solanum pennellii* introgression line population

ABSTRACT

Both rapid seed germination and seedling vigor are important for proper crop establishment and high yields. These traits appear to be controlled by several quantitative trait loci and are highly influenced by environmental conditions. In this study, we aimed to evaluate the seed germination performance and seedling vigor traits of several *Solanum pennellii* introgression lines (ILs) to select ideotypes for germination and vigor, and then investigate the selected ILs to identify candidate genes for these traits. We performed a germination test followed by seedling growth analysis in a collection of 46 different homozygous tomato ILs and their genitor accessions LA716 and cv. M-82, grown in the field for two consecutive growing seasons. As several traits were assessed, we used a multi-trait index based on factor analysis and genotype–ideotype distance (FAI-BLUP index) to rank the genotypes according to their seed germination performance and seedling vigor. After ranking, the introgressed segments of interest were selected for candidate gene identification using the map of the tomato IL population provided by the SOL Genomics Network. Using the FAI-BLUP index, IL 6-1, IL 4-1-1, IL 4-4, IL 11-4-1, IL 1-1-2, and IL 10-1 were selected as closest to the ideotype for seed germination performance and seedling vigor. Genes *At1g75350, At2g26350, At2g28490, At3g10920, At5g13200, At5g41480, and At4g21800* were identified as candidate genes for seed germination, while *At5g53000, At1g46480, At1g30755, At5g47390, At2g29630*, and *At4g03400* were identified as candidate genes for seedling vigor. Fine mapping, validation, and further investigation of these genes will provide valuable insights into the genetics of seed quality.

Keywords: *Solanum lycopersicum,* seed physiological quality, crop establishment

1. INTRODUCTION

The tomato (*Solanum lycopersicum* L.) is one of the most important vegetable crops worldwide owing to its high production (4.8 million hectares, with an average yield of 37 tons per hectare) (http://www.fao.org/faostat/en/; last accessed February 05, 2020). It is also an important source of vitamins and minerals in the human diet (Gürbüz Çolak et al., 2020). Tomato (*Lycopersicon esculentum* Mill) fruits can be consumed fresh or in other processed forms (e.g., paste, jam, etc.) and are among the most widely consumed vegetables worldwide. Globally, the tomato was the tenth most produced vegetable crop and the eleventh most produced crop in 2017 (Anon., 2019; Usanmaz and Atti, 2020; Halka et al., 2020). Tomato fruits are rich in vitamin C, minerals, and antioxidants (e.g., lycopene) and are known to be highly beneficial for human health (Jawad et al., 2020; Okatan, 2020). In addition to its high economic and nutritional value, the tomato is one of the best-studied fruit-bearing model organisms and has been used in several plant physiology, biochemistry, genetics, genomics, and breeding studies (Blanca et al., 2015; Fan et al., 2020; Kimura & Sinha, 2008).

To meet the increasing global demand for food, plant breeders have worked intensively to develop new tomato varieties with desirable flavor and texture traits and high yield potential (Schouten et al., 2019). According to Geshnizjani et al. (2020), improving seed quality is essential for sustainable and profitable crop production; therefore, it should be a target of tomato breeding programs. The ability to germinate quickly and uniformly and seedling vigor are essential seed quality attributes (Shrestha et al., 2016). These parameters may severely affect seedling establishment and the further growth of the plant. Thus, improving them ultimately improves crop production success (Finch-Savage & Bassel, 2016).

Seed germination performance and vigor are profoundly affected by the environmental conditions under which the seeds develop and mature, including temperature, humidity, light, and nutrients, as well as those that the seeds encounter after harvesting, such as drying, chemical treatments, and storage conditions (Kazmi et al., 2012). Regarding the genetic nature of these traits, studies have suggested that they are quantitative traits controlled by many genes (J.-P. Fernandez-Moreno et al., 2017; Hatzig et al., 2015).

An important determinant of seed quality and performance is the maternal environment in which the seeds develop and mature. During seed development, various environmental factors, such as temperature, light quality and quantity, and nutrients, may affect ultimate seed

quality. Therefore, seed quality is defined by both genetics (G) and the environment (E), as well as their interaction (G × E) (Koornneef, Bentsink, & Hilhorst, 2002; McDonald, 1998).

In tomatoes, as in many other crops, the domestication process has led to the loss of many potentially desirable traits (Doebley et al., 2006). For example, modern tomato cultivars are sensitive to non-optimal germination conditions, which limits their production to optimal environments (Foolad et al., 2003, 2007). The exploitation of natural variation to reintroduce lost genetic variation into cultivated species has been used in several species (Geshnizjani et al., 2020). To apply this approach, an interspecific IL population derived from a cross between *S. lycopersicum* cv. M-82 and *Solanum pennellii* Corr. was developed by Eshed and Zamir (1995). In this set of lines, the full genome of *S. pennellii* was represented as small introgressed fragments in 76 nearly isogenic lines (ILs) of the cultivar with an M-82 genetic background.

This population has been successfully used to identify the genomic regions associated with several traits of interest. More than 3000 QTL's associated with traits such as plant morphology and fruit metabolism have been identified using these ILs (Frary et al., 2010; Fridman et al., 2004; Lippman et al., 2007a; Rosental et al., 2016). Because *S. pennellii* is an abundant source of genetic variation, the *S. pennelli* IL population can be used to investigate and uncover candidate genes associated with seed germination performance and seedling vigor.

To uncover candidate genes for seed quality using the *S. pennellii* IL population, the first step is to select among the set of ILs that show good performance for this trait under different environments. Thus, at least two growing seasons should be considered because environmental conditions greatly influence seed performance. Seed germination and vigor can be assessed using several traits based on daily counting data from germination and seedling growth tests (L. J. da Silva et al., 2019). Selection indices are a suitable method of selecting the genotypes for these traits because selection is performed for more than one trait simultaneously, obtaining genotypes closer to an ideotype (Baker, 1986). According to Tandon and Jain (2004), the ideotype can be defined as an optimal combination of morphological and physiological traits that results in efficient matching of the plant material to its environment.

A selection index that combines factor analysis and genotype–ideotype design for multiple-trait selection—the FAI-BLUP index—was proposed by Rocha et al. (2018). It has been successfully used in the multi-trait selection of genotypes for several purposes (Oliveira et al., 2019; Rocha et al., 2019; M. J. Silva et al., 2018; Woyann et al., 2019). The advantages of using this index include that is gives the same weight to all traits, it simultaneously selects

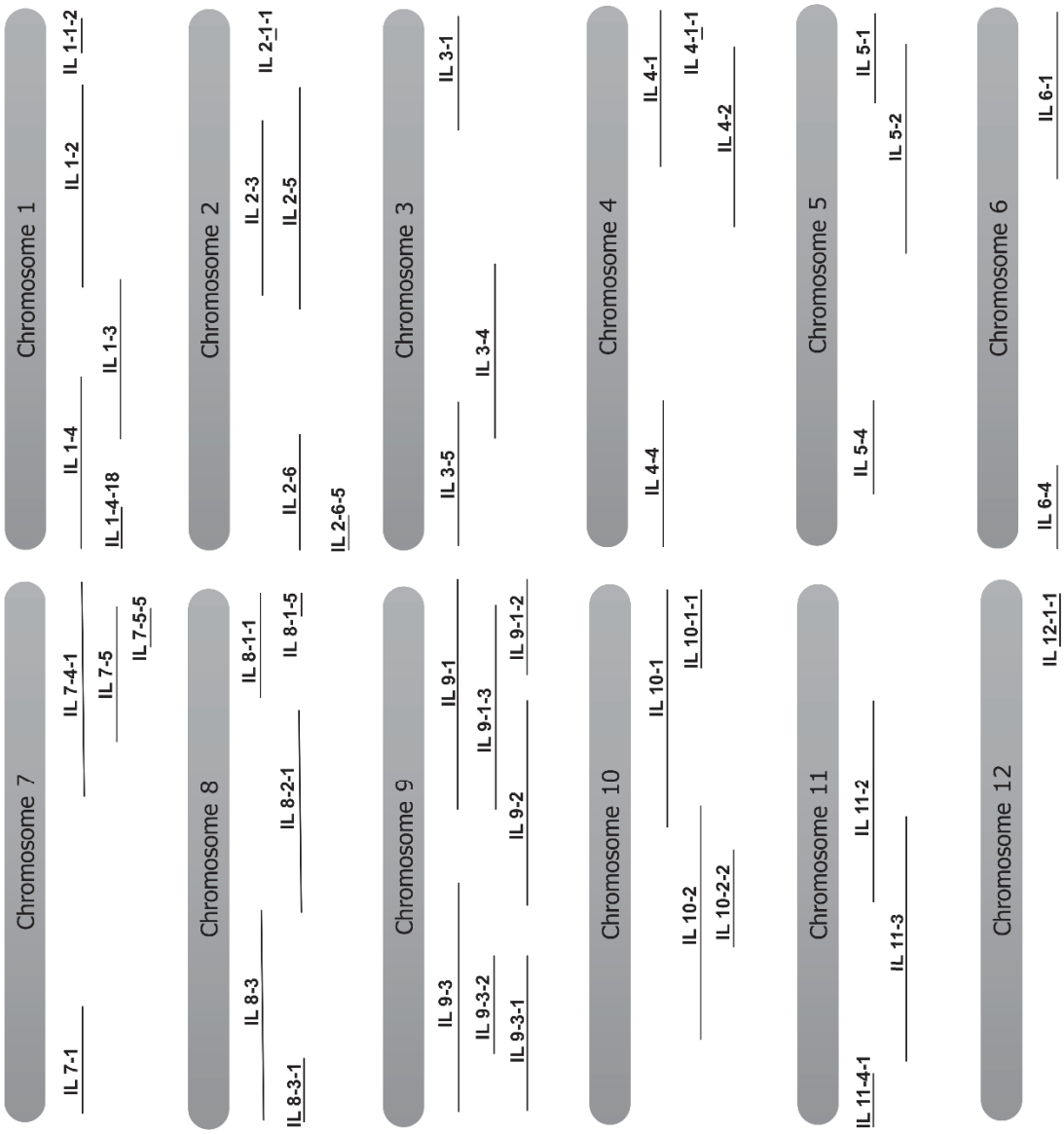
genotypes closer to the ideotype using a multi-trait approach free from multicollinearity, and it allows for the incorporation of multiple environments.

In this study, we evaluated the seed germination performance and seedling vigor traits of several *S. pennellii* ILs during two growing seasons to select genotypes that were similar to an ideotype using the FAI-BLUP index. Following selection, introgressed fragments from the selected ILs were investigated to identify candidate genes for these traits.

2. MATERIALS AND METHODS

2.1. ILS AND SEED PRODUCTION

For this study, we choose 46 ILs from the collection developed by Eshed and Zamir (1995), which covered all 12 tomato chromosomes (Supplementary Figure 1), and their genitors, cultivar M-82 and accession LA716.



Supplementary material 1. Scheme of the S. lycopersicum chromosomes showing the position of the introgression segments of the S. pennelli cromossome.

Seeds of the 48 genotypes were produced in two different growing seasons (2017 and 2018) in one of the experimental fields of the Agronomy Department at the Federal University of Viçosa, Brazil (20° 45′ 14″ S, 42° 52′ 53″ W). Seeds were collected from healthy mature fruits and subjected to natural fermentation for 2 days to remove the mucilaginous material surrounding the seed coat. After fermentation, the seeds were separated out using a fine-mesh sieve and washed with tap water to remove mucilage. The cleaned seeds were then dried on filter paper at room temperature and stored inside paper bags in a storage room (13 ºC and 30 % relative humidity) until use in subsequent experiments.

2.2. SEED GERMINATION PERFORMANCE AND SEEDLING VIGOR

This study was performed in vitro at the seed laboratory of the Federal University of Viçosa. In order to assess the environmental effects on seed quality, two experiments were carried out. For each experiment, seeds produced in the two different growing seasons were used. Seed germination and seedling vigor assessments were performed as follows.

Seeds of each genotype (four repetitions of 25 seeds) were placed in sterilized plastic plates with two layers of paper towels moistened with distilled water. The plates were incubated in a growth chamber at a constant temperature of 25 °C. Following incubation, the plates were examined daily for 14 days to determine the number of germinated seeds (normal seedlings). On the fourteenth day, when the experiment was complete, the seedling shoot and root lengths were measured manually.

The data from the daily seedling counts during the germination test and the seedling length measurements were processed using the software R and the package SeedCalc (Silva et al., 2019), to obtain the following germination traits: final germination percentage (FGP), germination index (GI), time to reach 10 % germination (T10), time to reach 50 % germination (T50), time to reach 90 % germination (T90), uniformity of germination (UnifG), mean germination time (MGT), variance of germination time (VarGer), coefficient of variation of germination time (CVT), germination synchrony (Sync), and uncertainty (Unc); and the following seedling traits: mean shoot (MSL), root (MRL), and total (MTL) lengths, uniformity index (UI), growth index (GrI), and vigor indices (VI1 and VI2). The equations used to calculate each trait, as described by Silva et al. (2019), are presented in Supplementary Table 1.

Supplementary material 2. Equations for obtaining the seed germination performance and vigour traits using data from daily counting of seedlingsseedling inon the germination/emergence test and seedling length measurement data.

Trait	Abbreviation	Description
Final germination percentage	FGP	FGP=(n/N)100 *n* is the number of germinated seeds, and *N* is the total number of seeds.
Germination index	GI	Calculated by the sum of the germinated seeds each day, divided by the number of days elapsed between sowing and germination.
Time to reach 10% germination	T_{10}	$$T_{10} = \frac{ti + \left\{\left[\frac{N}{\left(\frac{100}{10}\right)}\right] - ni\right\}(tj - ti)}{(nj - ni)}$$ N is the final number of seeds germinated, and n_i and n_j are the total number of seeds germinated in adjacent counts in time t_i and t_j, respectively, when $ni < \frac{N+1}{2} < nj$
Time to reach 50% germination	T_{50}	$$T_{50} = \frac{ti + \left\{\left[\frac{N}{\left(\frac{100}{50}\right)}\right] - ni\right\}(tj - ti)}{(nj - ni)}$$ N is the final number of seeds germinated, and n_i and n_j are the total number of seeds germinated in adjacent counts in time t_i and t_j, respectively, when $ni < \frac{N+1}{2} < nj$
Time to reach 90% germination	T_{90}	$$T_{90} = \frac{ti + \left\{\left[\frac{N}{\left(\frac{100}{90}\right)}\right] - ni\right\}(tj - ti)}{(nj - ni)}$$ N is the final number of seeds germinated, and n_i and n_j are the total number of seeds germinated in adjacent counts in time t_i and t_j, respectively, when $ni < \frac{N+1}{2} < nj$
Uniformity of germination	UnifG	$$UnifG = T_{90} - T_{10}$$ T90 is the time required for germination of 90% of the seeds, and T10 is the time required for germination of 10% of the seeds.
Mean Germination time	MGT	$$MGT = \sum_{i=1}^{K} n_i t_i \,/\, \sum_{i=1}^{K} n_i$$ *ni* is the number of seeds germinated per day (not the accumulated number, but the number

		corresponding to the *i-th* observation), and *ti* is the time since the beginning of the germination test up to the *i-th* observation.
Variance of the germination time	VarGer	$VarGer = \sum_{i=1}^{K} n(t_i - \bar{t})^2 \,/\, \sum_{i=1}^{K} (n_i - 1)$ $\bar{t}$ is the mean germination time, *ti* is the time between the beginning of the experiment and the *i-th* observation (day or hour), *ni* is the number of seeds germinated in the time *i*, and *k* is the last count of the germination test.
Coefficient of variation of germination time	CVT	$CVT = (St/\ \bar{t})100$ St is the standard deviation of the germination time, and $\bar{t}$ is the mean germination time.
Germination synchrony	Sync	$Sync = \sum Cni, 2/\ N$ Cni, 2 = ni (ni-1)/2 and $NN = \sum nntt(\sum nntt$ - 1)/2 *Cni* is the combination of the seeds germinated in the time *i*, two by two, and *ni* is the number of seeds germinated in the time *i*.
Uncertainty	Unc	$Unc = \sum_{i=1}^{k} fi\ log_2 fi$ With fi given by $fi = ni/\sum_{i=1}^{k} ni$ *fi* is the relative frequency of germination, and *ni* is the number of seeds germinated on the day *i*.
Mean shoot length	MSL	$MSL = \frac{\sum_{i=1}^{K} sl}{n}$ rl is the length of the root of each seedling, and n is the total number of seedlings evaluated
Mean root length	MRL	$MRL = \frac{\sum_{i=1}^{K} rl}{n}$ rl is the length of the root of each seedling, and n is the total number of seedlings evaluated
Mean total length	MTL	$MTL = \frac{\sum_{i=1}^{K} tl}{n}$ tl is the length of the root of each seedling, and n is the total number of seedlings evaluated
Uniformity index,	UI	

		$$UI = \left[1 - \frac{\sum_{i=1}^{n} \lvert Xi - X \rvert}{n\ x\ X}\right] x\ 100 - \left[n_{dead}\ x\ \frac{50}{n_{total}}\right]$$ *Xi* is the length of the seedling analyzed, *X* is the mean length of seedlings of the seed lot analyzed, *n* is the variable of total number of seedlings evaluated, *ndead* is the number of ungerminated or dead seeds present, and *ntotal* is the total number of seedlings.
Growth index	GrI	$$GI = [(mean(s))x\ (wh)] + [(mean(r)x\ (wr)]$$ mean(s) and mean(r) are the arithmetic means of shoot length and root length, respectively. ws and wr are adjustable weights in the formula for shoot and root, however, with reference values of 10 and 90, respectively
Vigor index 1	VI1	$$VI_1 = (Growth\ x\ wg) + (Growth\ x\ wu)$$ Growth is the growth index, and Uniformity is the uniformity index. wg and wu are adjustable weights in the formula for growth and uniformity, however, with standard values of 70 and 30, respectively
Vigor index 2	VI2	$$VI_2 = [(Growth\ x\ wg) + (Uniformity\ x\ wu)]\ x\ \left(\frac{G}{100}\right)$$

2.3. STATISTICAL ANALYSIS

The germination and seedling growth data from the two growing seasons were analyzed using mixed-model methodology via REML/BLUP (restricted residual maximum likelihood/best linear unbiased prediction), according to Patterson and Thompson (1971) and Henderson (1975), in the R software package lme4.

The statistical model was denoted by:

$$y = Xm + Zg + Wb + Ti + \varepsilon$$

where y is the data vector; m is the vector of the effects of the measurement-replication combination (assumed as fixed) added to the overall mean; g is the vector of genotype effects (assumed as random); b is the vector of environmental effects (assumed as random); i is the vector of the genotype × environment effects; ε is the vector of residue (random); *and* X, Z, W, and T represent the incidence matrices for these effects.

For the random effects of the model, the significance of the likelihood ratio test was assessed using the chi-square statistic with one degree of freedom. Genetic values (BLUP means) were predicted for each of the 48 genotypes based on the 18 evaluated traits over two growing seasons.

2.3.1. ILS RANKING

The genetic values (BLUP means) were submitted to a multi-trait index based on factor analysis and genotype–ideotype design, known as the FAI-BLUP index, to rank the genotypes. Principal component analysis, factor analysis, ideotype determination, and genotype–ideotype distance were performed in R software using the FAI-BLUP index routine developed by Rocha et al. (2018).

Principal component analysis was used to extract the factorial loads of the genetic correlation matrix obtained from the genetic values. The varimax criterion described by Kaiser (1958) was used for analytic rotation, and the weighted least squares method described by Bartlett (1938) was used to calculate the factor scores.

The number of ideotypes was defined based on a combination of desirable and undesirable selection factors. The following algorithm provides the number of ideotypes:

$$NI = 2^n,$$

in which NI is the number of ideotypes and n is the number of factors.

The ideotypes for seed germination performance and seedling vigor were determined by considering ideal values for each trait (the minimum or maximum values of the dataset). The ideotype was built considering the maximum predicted genetic values for the traits FGP, GI, UnifG, Sync, MSL, MRL, MTL, UI, GrI, VI1, and VI2, and the minimum predicted genetic values for the traits T10, T50, T90, MGT, Unc, VarGer, and CVT.

After ideotype determination, genotype–ideotype distances were estimated and converted into spatial probabilities, enabling genotype ranking. The following algorithm was used:

$$P_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{i=1;j=1}^{i=n;j=m} \frac{1}{d_{ij}}}$$

Where, Pij is the probability that the i^{th} genotype (i = 1, 2,..., n) is similar to the j^{th} ideotype (j = 1, 2,..., m) and d_{ij} is the genotype–ideotype distance from the i^{th} genotype to the j^{th} ideotype, based on the standardized mean Euclidean distance.

2.4. CANDIDATE GENE IDENTIFICATION

The genomic segments of *S. pennellii* introgressed into the selected ILs were further analyzed to identify candidate genes for seed germination performance and seedling vigor, adapting the methodology described previously (Toubiana et al., 2012). We used a map of the tomato IL population provided by the SOL Genomics Network (http://solgenomics.net/). This map displays individual chromosomes with restriction sites for the different introgressed segments. It also displays all identified marker genes.

We adapted the script developed in Toubiana et al. (2012) to automatically identify all marker genes from each selected IL in the HTML code. The functionality of each identified gene was inferred manually using information on the respective orthologs in the *Arabidopsis thaliana* genome from the TAIR database (https://www.arabidopsis.org/). The identification of each candidate gene for germination performance and seedling vigor was based on the relevance of its functionality to these processes.

3. RESULTS

The effects of the treatments (ILs and genitors), environment (growing seasons 1 and 2), and treatment × environment interactions were significant ($p < 0.01$) for all the evaluated traits. As all traits showed significant treatment effects, they were all included in the FAI-BLUP analysis.

By analyzing the eigenvalues and cumulative frequency for the principal components obtained by the genetic correlation matrix, we found that the first five components presented eigenvalues > 1 (Table 1). Thus, according to Kaiser's criterion (Kaiser, 1958), the data could be dimensionally reduced to five factors. The cumulative variance of the first five principal components accounted for ~92 % of all the genetic variability present in the dataset.

Table 1. Eigenvalue estimates from principal components analysis and the proportion of variance explained by them

Principal component	Eigenvalue	Cumulative variance
PC1	8.743	48.573
PC2	2.923	64.814
PC3	2.437	78.356
PC4	1.427	86.283
PC5	1.010	91.896
PC6	0.450	94.393
PC7	0.405	96.642
PC8	0.270	98.141
PC9	0.149	98.970
PC10	0.077	99.399
PC11	0.048	99.666
PC12	0.035	99.863
PC13	0.016	99.949
PC14	0.007	99.988
PC15	0.002	100.000
PC16	2.80E-05	100.000
PC17	1.49E-06	100.000
PC18	1.69E-08	100.000

The clustering of the traits in each factor is shown in Figure 1, and the factorial loadings after varimax rotation of the five factors are shown in Figure 2. In the factor analysis, traits clustered within one factor showed higher correlations with each other than with traits in other factors. In addition, the genetic correlations between the traits within each factor can be in the same and/or opposite directions. In this analysis, the first factor showed strong genetic

correlations with MSL, MRL, MTL, GrI, and VI1. The second factor showed strong genetic correlations with T10, T50, and MGT. For the third factor, T90, UG, VarGer, CVT, Sync, and Unc exhibited the highest correlation values. For the fourth factor, FGP, GI, and VI2 showed the highest correlation values. Finally, the fifth factor included only UnifG.

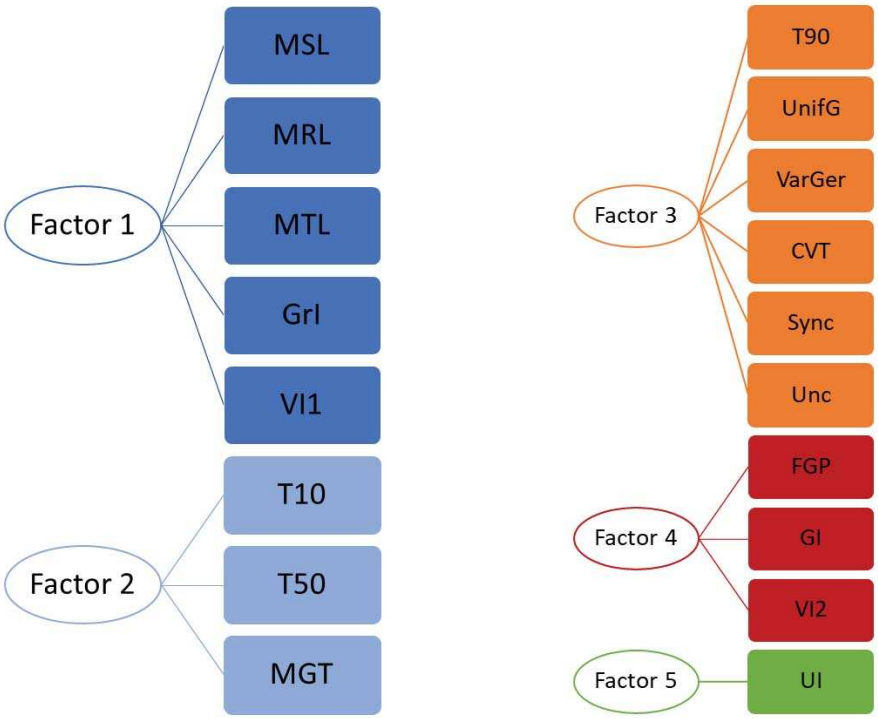


Figure 1. Traits evaluated in the seeds of 48 tomato genotypes clustered in five factors: seedling size, germination time, vigor, germination, and uniformity. FGP = final germination percentage, GI = germination index, T10 = time to reach 10 % germination, T50 = time to reach 50 % germination, T90 = time to reach 90 % germination, UnifG = uniformity of germination, MGT = mean germination time, VarGer = variance of germination time, CVT = coefficient of variation of germination time, Sync = synchronization of germination, Unc = uncertainty, MSL = mean shoot length, MRL = mean root length, MTL = mean total length, UI = uniformity index, GrI = growth index, VI1 = vigor index 1, VI2 = vigor index 2.

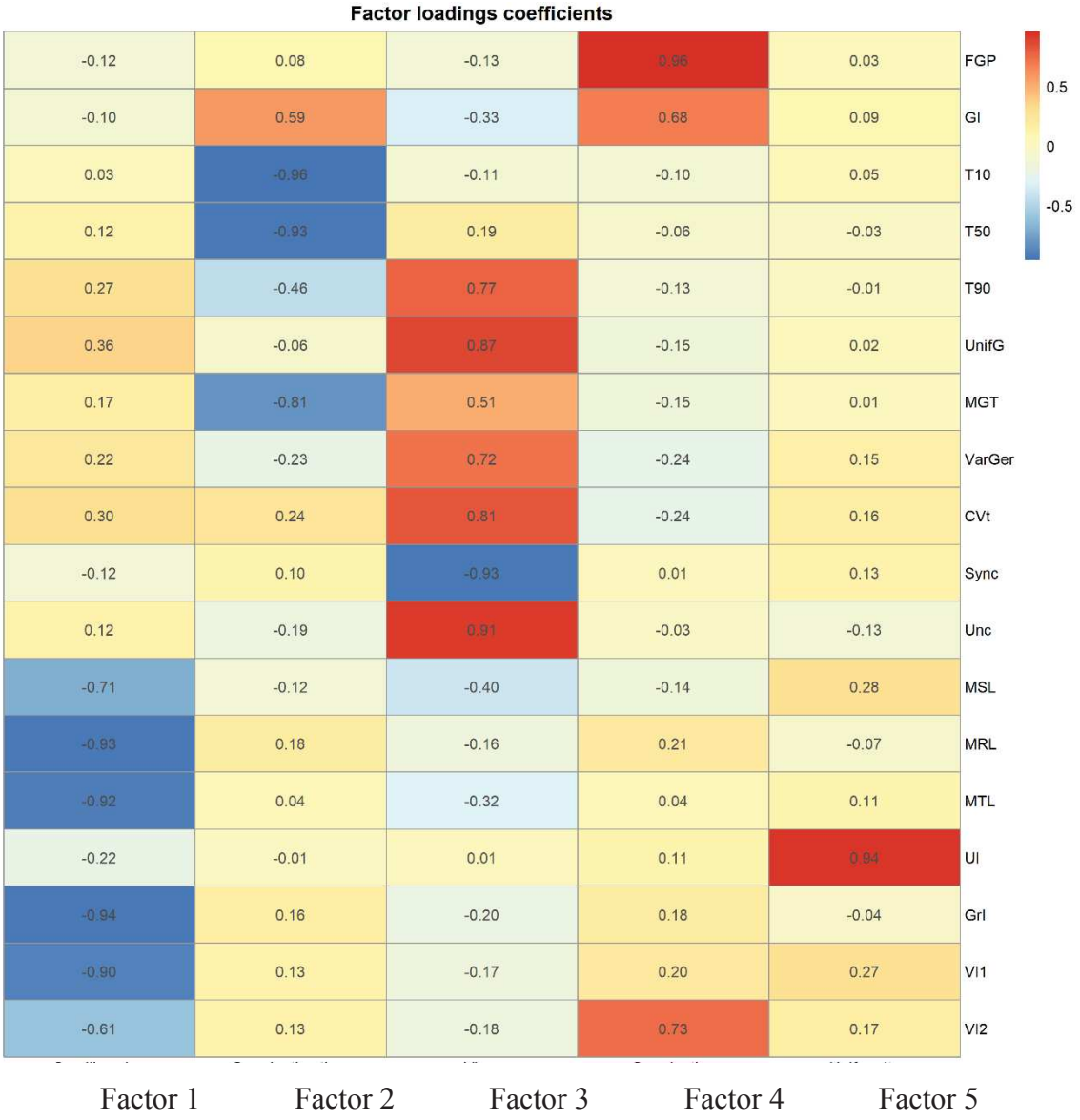


Figure 2. Heat map showing the factorial loadings after varimax rotation for five factors: seedling size, germination time, vigor, germination, and uniformity. The color indicates the strength and direction of correlations among traits in a factor; bluer is more negative, redder is more positive. FGP = final germination percentage, GI = germination index, T10 = time to reach 10 % germination, T50 = time to reach 50 % germination, T90 = time to reach 90 % germination, UnifG = uniformity of germination, MGT = mean germination time, VarGer = variance of germination time, CVT = coefficient of variation of germination time, Sync = synchronization of germination, Unc = uncertainty, MSL = mean shoot length, MRL = mean root length, MTL = mean total length, UI = uniformity index, GrI = growth index, VI1 = vigor index 1, VI2 = vigor index 2.

The ideotype for seed germination performance and seedling vigor was designed so that it showed desirable traits for these five factors. Figure 3 shows the ranking of the 48 genotypes according to the FAI-BLUP index and the probability, in relation to ideotype distance, of good seed germination performance and seedling vigor. The cultivar M-82 was ranked seventh closest to the ideotype, and the accession LA716 was ranked as the most distant genotype. The most vigorous genotypes, ranked according to the FAI-blup index, were IL 11-4-1, IL 4-4, IL 1-1-2, and IL 6-1. Therefore, we selected these genotypes to investigate the candidate genes.

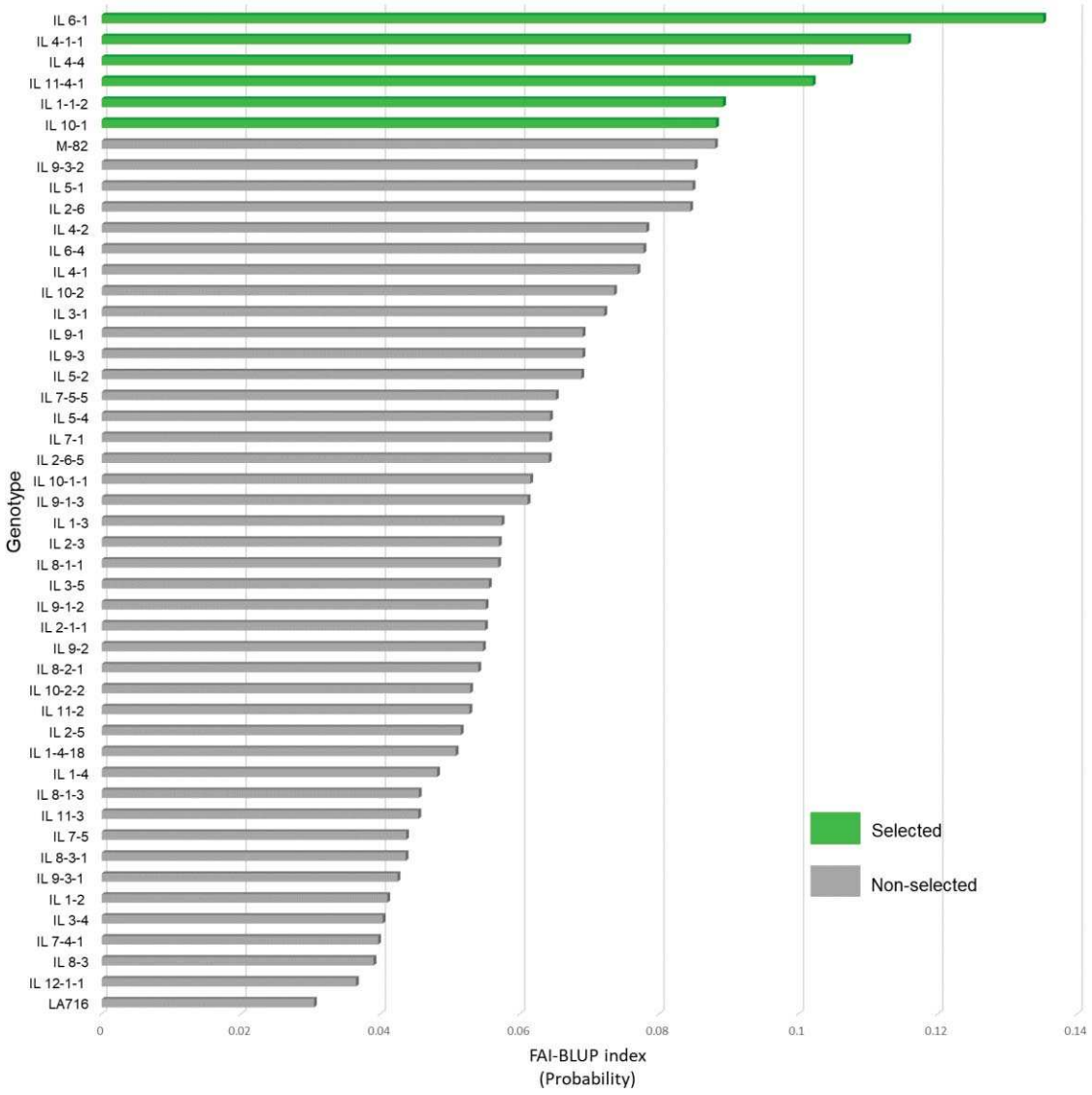


Figure 3. Genotype ranking and selected genotypes using the FAI-BLUP index.

The BLUP means for each genotype and trait were used to compare the performance of each of the evaluated genotypes to that of M-82 via a colored gradient (Figure 4). The genotypes were ranked for each trait individually according to the proposed ideotype: maximum values

for FGP, GI, UnifG, Sync, MSL, MRL, MTL, UI, GrI, VI1, and VI2 and minimum values for T10, T50, T90, MGT, Unc, VarGer, and CVT.

Introgression of the S. *pennelli* genome into M-82 promoted visible changes in all evaluated traits for most ILs. Introgression can lead both increases and decreases in these traits (Figure 4). The individual ranking for each trait showed coherence with the multi-trait ranking; LA716 and IL6-1 were ranked furthest from and closest to the ideotype, respectively, according to the FAI-BLUP index (Figure 3) and were ranked as the worst and best genotypes, respectively, for most individual traits (Figure 4).

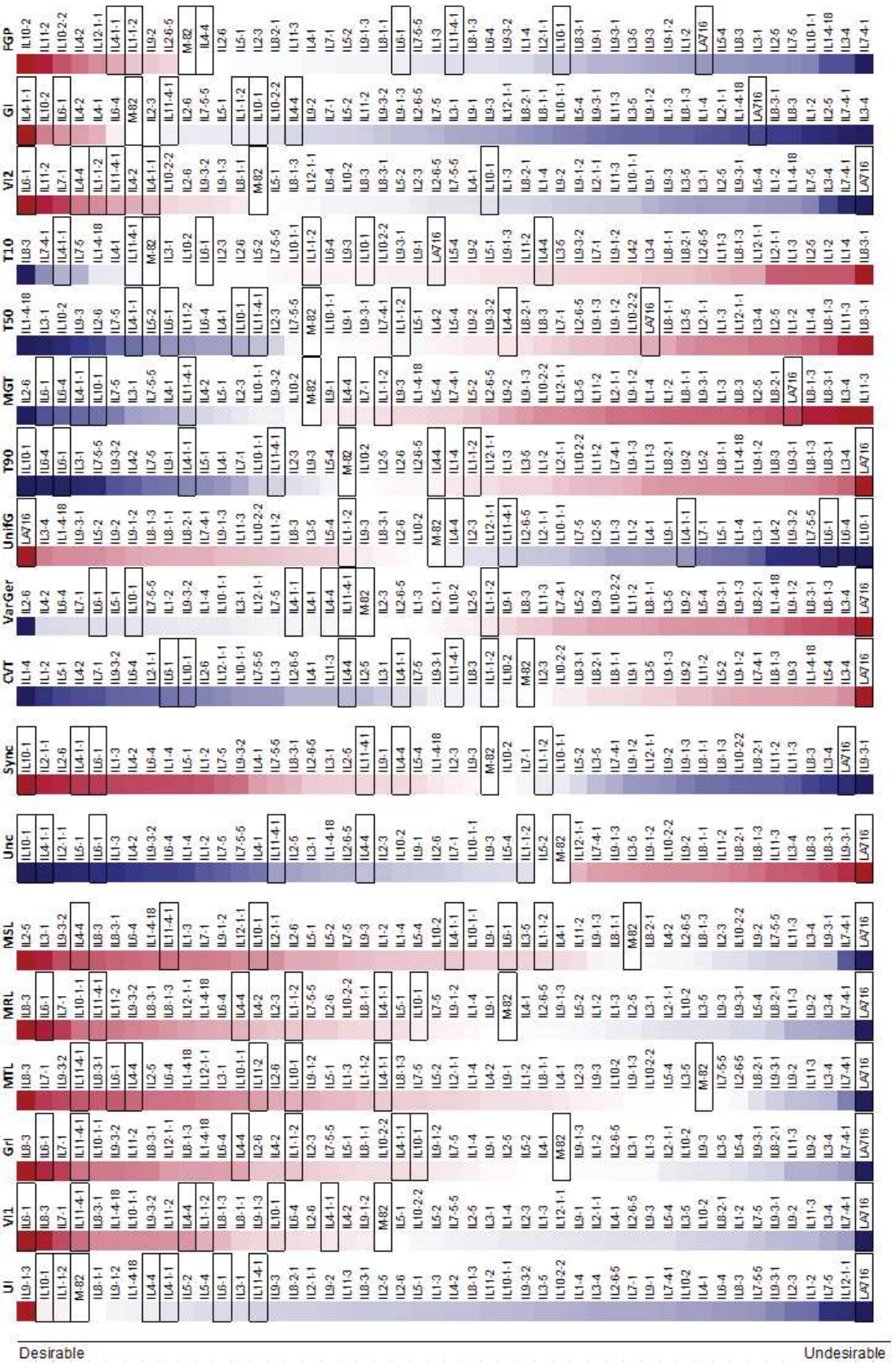


Figure 4. Heat map of BLUP means of the evaluated traits. Genotypes were compared to M-82 via a colored gradient, the intensity of which ranges from dark blue, corresponding to the

most significant reductions in the traits, to dark red, corresponding to the highest increases in the traits. The genotypes were ranked for each trait according to the ideotype; on the left are the closest genotypes to the proposed ideotype for each trait. The ILs selected using the FAI-BLUP index, M-82, and LA176 are highlighted with rectangles. FGP = final germination percentage, GI = germination index, T10 = time to reach 10 % germination, T50 = time to reach 50 % germination, T90 = time to reach 90 % germination, UnifG = uniformity of germination, MGT = mean germination time, VarGer = variance of germination time, CVT = coefficient of variation of germination time, Sync = synchronization of germination, Unc = uncertainty, MSL = mean shoot length, MRL = mean root length, MTL = mean total length, UI = uniformity index, GrI = growth index, VI1 = vigor index 1, VI2 = vigor index 2.

The genotypes IL 6-1, IL 11-4-1, IL 4-4, IL 1-1-2, and IL 10-1, selected using the FAI-BLUP index, were further analyzed for candidate genes. The gene markers associated with the introgressed segments from the selected ILs were identified via the SOL Genomics Network, and their functionality was inferred using information on their respective orthologs in the *A. thaliana* genome.

No genes related to germination or seedling vigor were identified in the selected IL 4-1-1. In IL 6-1, the genes At3g10920, At4g21800, At5g13200, and At5g41480 were identified as candidate genes for seed germination performance (Figure 5). At3g10920 encodes manganese superoxide dismutase (MSD1) and is involved in embryonic development and seed dormancy (Irani and Todd, 2018). At4g21800 encodes QQT2, which is required for early embryonic development and participates in the organization of microtubules during cell division (Kwak et al., 2017). At5g13200 encodes a protein involved in the abscisic acid (ABA)-mediated regulation of seed germination and dormancy (Liu et al., 2013). At5g41480 encodes dihydrofolate synthetase, which is involved in folate biosynthesis, embryo development, and ending seed dormancy. In IL 4-4, the genes At1g75350 and At2g26350 were identified (Figure 5). At1g75350 encodes the protein L31 (Bryant et al. 2011) and At2g26350 encodes the zinc-binding peroxisomal integral membrane protein PEX10 (Bartel et al., 2014). These two proteins are required for embryogenesis and are involved in embryonic development. Finally, in IL 11-4-1, the gene At2g28490 was identified. It encodes RmlC-like cupin superfamily proteins and has a molecular function in nutrient reservoir activity; therefore, it is related to germination performance (Yobi et al., 2020).

The candidate genes for seedling vigor in IL 6-1 were At5g47390 and At2g29630. At5g47390 encodes a transcription factor that specifically controls cell expansion and is

involved in auxin homeostasis, lateral root development, leaf development, the positive regulation of cell growth, and post-embryonic plant organ morphogenesis (Lu et al., 2014). At2g29630 encodes a protein involved in thiamin biosynthesis (Kong et al., 2008). In IL 4-4, At1g46480 and At1g30755 were identified (Figure 5). At1g46480 encodes WOX4, a WUSCHEL-related homeobox gene family involved in cell division, procambium histogenesis, and tissue and vasculature development (Agusti et al., 2011; Nic-Can et al. 2013). At1g30755 encodes an elongation factor (DUF668) involved in the positive regulation of seedling root and shoot growth (Stührwohldt et al., 2014). In IL 1-1-2, At5g53000 was identified (Figure 5), which encodes the Tap46 regulatory subunit of protein phosphatase 2A and is related to the positive regulation of cell growth (Kataya et al., 2016). Finally, in IL 10-1, the gene At4g03400 was identified (Figure 5), which encodes a GH3-related gene involved in red light-specific hypocotyl elongation.

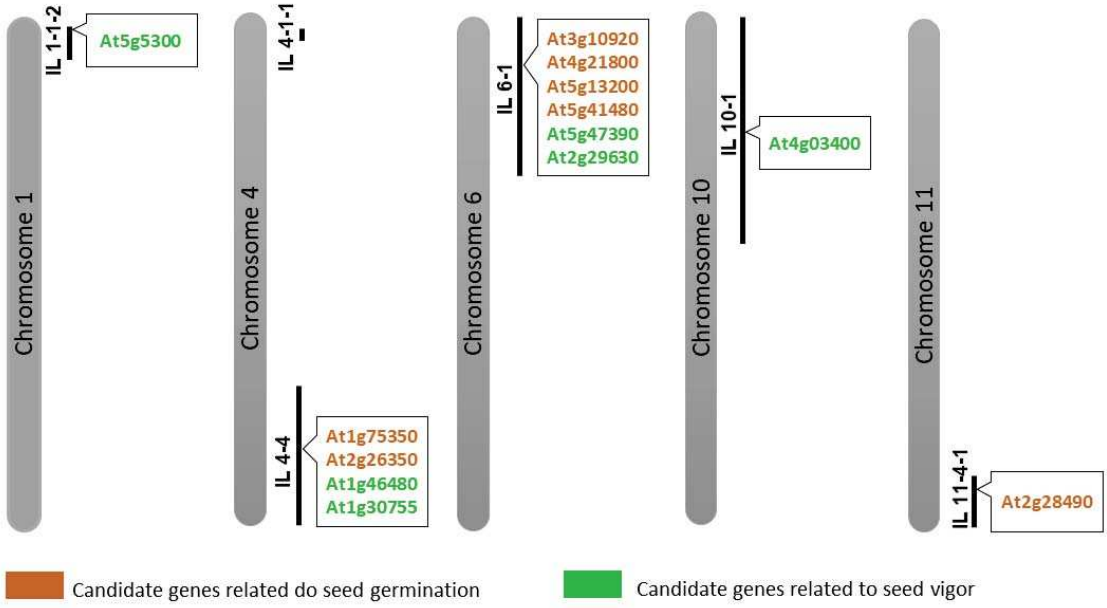


Figure 5. Candidate genes related to seed germination performance and seedling vigor.

4. DISCUSSION

The significant effects of treatment on all the evaluated traits demonstrates the wide variability for those traits among the genotypes. In our study, the genetic variability among the genotypes was caused by the small genomic segments of *S. pennellii* that replaced the homologous regions in the M-82 background (Eshed & Zamir, 1995b). Genetic variability among genotypes is essential for genetic progress (Hmielowski, 2018).

The significant effect of the environment (growth season) on all the evaluated traits indicates that variations in environmental conditions during seed development influence seed formation and, subsequently, the expression of genes involved in the control of the evaluated traits. It is well known that environmental conditions such as soil fertility, water supply, and temperature have a significant influence on seed development (Rosental et al., 2016). Thus, seeds produced in different growing seasons, which develop under slightly different environmental conditions, will have some differences in their physiological quality. In addition, the treatment × environment interaction had a significant effect ($p < 0.01$) on all traits, indicating the differential behavior of genotypes in different environments.

The FAI-BLUP selection index allows progenies to be ranked while considering multiple traits and environments simultaneously. Using the genetic values obtained from individual analyses, the index capitalizes on the effects of genotype × growth season interactions. According to Rocha et al. (2018), the FAI-BLUP index is based on factors, whereby traits clustered within the same factor have high correlations and those in different factors have low correlations. Thus, selection toward the ideotype, considering the information of each environment, enhances the effects of genotypes × growth season interactions. In addition, this index is free from multicollinearity, which is a significant advantage as not all the traits evaluated in this study were orthogonal.

LA716 was ranked as the genotype furthest from the seed germination performance and seedling vigor ideotype (Figure 3). When evaluating the ranking for each trait separately (Figure 4), LA716 was among the most distant genotypes from the ideotype for almost all traits. For traits related to seedling size, the explanation for this result is the size of the *S. pennelli* seeds. The seeds of this wild tomato are several times smaller than the average size of cultivated tomato seeds (Doğanlar et al., 2000). According to Khan et al. (2012), initial seedling size is positively related to seed size, and larger seeds have a better seedling survival rate and higher competitiveness both within and among species. Therefore, the smaller seeds of LA716

accounted for its poor early seedling development performance. As for the other traits, it is important to point out that in our study, the accession LA716, a wild tomato species, was compared to genotypes with all or almost all the genetic background of the modern commercial tomato cultivar M-82. The agronomic traits of interest in this cultivar have been selected for and improved during domestication and breeding (Doebley et al., 2006). Therefore, although using wild species to recover some desirable traits lost during these processes is a valuable approach in modern breeding, the performance of wild species *per se* is generally worse than that of commercial cultivars.

Cultivar M-82 was ranked seventh closest to the seed germination performance and seedling vigor ideotype. This indicates that this genotype already produces very vigorous seeds, which is expected because M-82 is a commercial cultivar. The introgression of *S. pennelli* chromosome fragments into M-82 affected the performance of all the ILs in most traits, as shown in Figure 4. This introgression can account for either an increase or a decrease in a trait and can also promote increases in one trait and decreases in another (Figure 4). According to Kazmi et al. (2017), seed germination performance and seedling vigor are quantitative traits that are controlled by multiple genes. The findings of this study add to the knowledge that the changes in performance observed in this set of evaluated ILs are the result of interactions between the introgressed *S. pennellii* genomic fragments and the rest of the M-82 genome.

IL 6-1 was ranked closest to the seed germination and seedling vigor ideotypes (Figure 3). It also ranked among the best genotypes for almost all the evaluated traits (Figure 4). In this IL, four genes related to seed germination—*At3g10920*, *At4g21800, At5g13200* and *At5g41480*—and two genes related to seedling vigor—*At5g47390* and *At4g29630*—were identified (Figure 5).

At3g10920 encodes MSD1 and is involved in embryo development and ending seed dormancy (Irani & Todd, 2018). The superoxide dismutase (SOD) is one of the major reactive oxygen species (ROS) scavenging enzymes of plants. SODs catalyze the dismutation of $O2^{-}$ to H_2O_2, constituting a front-line defense against ROS. In an experiment overexpressing MSD1 (Xi et al., 2010), single, double, and triple MSD1-transgene overexpressers displayed remarkably enhanced oxidative stress tolerance during seed germination and early seedling growth. In addition, an increase in total catalase activity was observed in the single MSD1-transgenic lines as a result of MSD1 overexpression. The authors concluded that the combined increase in Mn-SOD and CAT activities in the seeds played an essential role in the improvement of antioxidant capacity at an early developmental stage in *A. thaliana*. A good antioxidant

response significantly improves germination success; therefore, *At3g10920* may account for the outstanding performance of IL6-1 in most of the traits associated with germination performance.

Another gene that might account for this high performance is *At5g13200*. *At4g21800* encodes QQT2, which is required for early embryo development and participates in the organization of microtubules during cell division (Kwak et al., 2017). QQT proteins co-localize with microtubules and are essential for early embryonic development (Lahmy et al., 2007). Seeds in which the embryos are not well-developed are much less effective in germinating normal seedlings. *At5g13200* encodes a protein with an unknown function that is involved in the ABA-mediated regulation of seed germination and dormancy (Liu et al., 2013). ABA regulates many aspects of plant growth and development, including embryo maturation, seed dormancy, germination, cell division and elongation, and responses to environmental stresses such as drought, salinity, cold, pathogen attack, and UV radiation (Finkelstein & Rock, 2002). *At5g41480* encodes dihydrofolate synthetase. This protein is involved in folate biosynthesis, embryonic development, and ending seed dormancy (Wang et al., 2020). Thus, these genes may play essential roles in the excellent germination performance of IL-6-1.

As for the traits related to seedling vigor, IL 6-1 showed better performance than M-82 for all traits, and it was ranked among the best genotypes for most of them (Figure 4). The At5g47390 and At2g29630 genes were identified as candidate genes for these traits in IL 6-1. *At5g47390* encodes a transcription factor that specifically controls cell expansion and is involved in auxin homeostasis, lateral root development, leaf development, the positive regulation of cell growth, and post-embryonic plant organ morphogenesis (Lu et al., 2014). *At2g29630* encodes a protein involved in thiamin biosynthesis (Kong et al., 2008). A severe reduction in THIC levels in plants decreases vitamin B1 (thiamine diphosphate) levels, and also leads to changes in the levels of numerous other metabolites because many primary metabolic enzymes require a thiamine diphosphate co-factor. ThiC mutants were chlorotic and arrested at the cotyledon stage (Przybyla-Toscano et al., 2018). Therefore, we hypothesize that the presence of *At5g47390* and *At2g29630* introgressed from *S. pennelli* in IL 6-1 can account for the outstanding vigor shown by the seedlings.

IL 4-1-1 was selected as the second genotype closest to the seed germination performance and seedling vigor ideotype (Figure 3); however, we did not identify candidate genes that could justify this result. In our approach, genes with unknown functionalities, or those that do not have an *A. thaliana* ortholog, were not included in the list of potential

candidate genes; therefore, it is possible that one or several of the excluded genes are related to the seed germination performance and seedling vigor observed for IL 4-1-1 in this study.

In IL 4-4, the orthologs *At1g75350* that encodes protein L31 (Bryant et al., 2011) and *At2g26350* that encodes a zinc-binding peroxisomal integral membrane protein PEX10 (Bartel et al., 2014) were identified. These two proteins are required for embryogenesis and are involved in embryonic development. PEX10 knockdown mutants produce only a small number of viable seeds and most of the embryos show immature and lethal phenotypes (Nito et al., 2007). Thus, the presence of these genes could be associated with the improvement in the germination performance of IL 4-4 in relation to M-82. The introgression of the *S. pennellii* genome carrying these genes increased the performance of several traits (VarGer, CVT, Sync, Unc, MSL, MRL, MTL, GrI, VI, and UI) (Figure 4).

In addition, the genes *At1g46480* and *At1g30755,* which are related to seedling vigor, were also identified in IL 4-4. *At1g46480* encodes WOX4, which is a WUSCHEL-related homeobox gene family (Agusti et al., 2011). According to Nic-Can et al. (2013) WOX4 regulates plant vascular proliferation and is involved in cell division, procambium histogenesis, and tissue and vasculature development. *At1g30755* encodes an elongation factor (DUF668) involved in the positive regulation of seedling root and shoot growth (Stührwohldt et al., 2014), which can explain the fact that IL 4-4 was ranked among the best genotypes for traits related to seedling growth (Figure 4).

Analysis of the performance of IL 1-1-2 in the seedling growth parameters MSL, MRL, and MTL (Figure 4) showed that it was ranked closer to the ideotype than M-82, indicating that the chromosomic fragment of *S. pennelli* contained genes that, either individually or together with the genes of M-82, enhanced its performance. We identified *At5g53000* in this IL. This gene encodes the Tap46 regulatory subunit of protein phosphatase 2A and is involved in the positive regulation of cell growth (Kataya et al., 2016). According to Bheri and Pandey (2019) phosphatases act as a "molecular switch," which turns the regulatory processes driving growth and development on or off under normal circumstances. Thus, the outstanding seedling growth of this IL might be associated with the gene *At5g53000*.

The genotype M-82 was among the 10 worst genotypes for the seedling growth trait MSL; IL 10-1 was much better ranked. This increase in performance was caused by introgression of the *S. pennelli* genome. Analyzing the genes present in the introgressed portion in IL 10-1, the ortholog *At4g03400*, which encodes a GH3-related gene involved in red light-specific hypocotyl elongation, was identified as a candidate gene for seedling vigor (Nakata et

al., 2018). A previous analysis performed by Miyazaki et al. (2016) suggested that this gene is involved in the determination of hypocotyl elongation. Therefore, we can assume that *At4g03400,* a gene inherited from *S. pennellii*, was associated with the increase in the growth performance of this IL.

Consistent with the proposition that seed germination performance and seedling vigor are quantitative traits controlled by multiple genes (Kazmi et al., 2017), we identified several candidate genes that might be related to them. Fine mapping, validation, and further investigation of these genes will provide valuable insights into the genetics of tomato seed germination performance and seedling vigor. Furthermore, our results represent a starting point for future studies to advance our understanding of the molecular and genetic basis underlying seed germination performance and seedling vigor.

5. CONCLUSION

The FAI-BLUP index ranked IL 6-1, IL 4-1-1, IL 4-4, IL 11-4-1, IL 1-1-2, and IL 10-1 as closest to the ideotype of seed germination performance and seedling vigor. The genes *At1g75350, At2g26350, At2g28490, At3g10920, At5g13200, At5g41480,* and At4g21800 were identified as candidate genes for seed germination. The genes *At5g53000, At1g46480, At1g30755, At5g47390*, *At2g29630*, and *At4g03400* were identified as candidate genes for seedling vigor.

REFERENCES

Agusti, J., Lichtenberger, R., Schwarz, M., Nehlin, L., Greb, T., 2011. Characterization of transcriptome remodeling during cambium formation identifies MOL1 and RUL1 as opposing regulators of secondary growth. PLoS Genet. 7. https://doi.org/10.1371/journal.pgen.1001312

Baker, R. J. (2020). Selection indices in plant breeding. CRC Press.

Bartlett, M. S., 1937. The statistical conception of mental factors. British Journal of Psychology, 28, 97–104. https://doi.org/10.1111/j.2044-8295.1937.tb00863.x.

Bartel, B., Burkhart, S.E., Fleming, W.A., 2014. Protein transport in and out of plant peroxisomes, in: Molecular Machines Involved in Peroxisome Biogenesis and Maintenance. Springer-Verlag Wien, pp. 325–345. https://doi.org/10.1007/978-3-7091-1788-0_14

Bernardo, S., Soares, A. A., & Mantovani, E. C. 2019. Manual de irrigação. (Editora UFV).

Bheri, M., Pandey, G.K., 2019. PP2A Phosphatases Take a Giant Leap in the Post-Genomics Era. Curr. Genomics 20, 154–171. https://doi.org/10.2174/1389202920666190517110605

Blanca, J., Montero-Pau, J., Sauvage, C., Bauchet, G., Illa, E., Díez, M.J., Francis, D., Causse, M., van der Knaap, E., Cañizares, J., 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. BMC Genomics 16, 257. https://doi.org/10.1186/s12864-015-1444-1

Bryant, N., Lloyd, J., Sweeney, C., Myouga, F., Meinke, D., 2011. Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in Arabidopsis. Plant Physiol. 155, 1678–1689. https://doi.org/10.1104/pp.110.168120

Doebley, J.F., Gaut, B.S., Smith, B.D., 2006. The Molecular Genetics of Crop Domestication. Cell. https://doi.org/10.1016/j.cell.2006.12.006

Doganlar, S., Frary, A., Tanksley, S.D., 2000. The genetic basis of seed-weight variation: Tomato as a model system. Theor. Appl. Genet. 100, 1267–1273. https://doi.org/10.1007/s001220051433

Eshed, Y., Zamir, D., 1995. An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141.

Fan, Y., Yang, W., Yan, Q., Chen, C., Li, J., 2020. Genome-wide identification and expression analysis of the protease inhibitor gene families in tomato. Genes (Basel). 11. https://doi.org/10.3390/genes11010001

FAO (2021). Food and Agriculture Organization of the Union. https://doi.org/10.4060/cb2395fr (Consulted on 10/02/2022).

Fernandez-Moreno, J.-P., Levy-Samoha, D., Malitsky, S., Monforte, A.J., Orzaez, D., Aharoni, A., Granell, A., 2017. Uncovering tomato quantitative trait loci and candidate genes for fruit cuticular lipid composition using the Solanum pennellii introgression line population. J. Exp. Bot. 68, 2703–2716. https://doi.org/10.1093/jxb/erx134

Finch-Savage, W.E., Bassel, G.W., 2016. Seed vigour and crop establishment: Extending performance beyond adaptation. J. Exp. Bot. https://doi.org/10.1093/jxb/erv490

Finkelstein, R.R., Rock, C.D., 2002. Abscisic Acid Biosynthesis and Response. Arab. B. 1, e0058. https://doi.org/10.1199/tab.0058

Foolad, M.R., Subbiah, P., Zhang, L., 2007. Common QTL affect the rate of tomato seed germination under different stress and nonstress conditions. Int. J. Plant Genomics 2007. https://doi.org/10.1155/2007/97386

Foolad, M.R., Zhang, L.P., Subbiah, P., 2003. Genetics of drought tolerance during seed germination in tomato: inheritance and QTL mapping. Genome 46, 536–545. https://doi.org/10.1139/g03-035

Frary, A., Göl, D., Keleş, D., Ökmen, B., Pınar, H., Şığva, H.Ö., Yemenicioğlu, A., Doğanlar, S., 2010. Salt tolerance in Solanum pennellii: antioxidant response and related QTL. BMC Plant Biol. 10, 58. https://doi.org/10.1186/1471-2229-10-58

Fridman, E., Carrari, F., Liu, Y.S., Fernie, A.R., Zamir, D., 2004. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science (80-.). 305, 1786–1789. https://doi.org/10.1126/science.1101666

Geshnizjani, N., Snoek, B.L., Willems, L.A.J., Rienstra, J.A., Nijveen, H., Hilhorst, H.W.M., Ligterink, W., 2020. Detection of QTLs for genotype × environment interactions in tomato seeds and seedlings. Plant Cell Environ. 43, 1973–1988. https://doi.org/10.1111/pce.13788

Gürbüz Çolak, N., Eken, N.T., Ülger, M., Frary, A., Doğanlar, S., 2020. Mapping of quantitative trait loci for antioxidant molecules in tomato fruit: Carotenoids, vitamins C and E, glutathione and phenolic acids. Plant Sci. 292, 110393. https://doi.org/10.1016/j.plantsci.2019.110393

Hatzig, S. V., Frisch, M., Breuer, F., Nesi, N., Ducournau, S., Wagner, M.H., Leckband, G., Abbadi, A., Snowdon, R.J., 2015. Genome-wide association mapping unravels the genetic control of seed germination and vigor in Brassica napus. Front. Plant Sci. 6. https://doi.org/10.3389/fpls.2015.00221

Henderson, C.R., 1975. Best Linear Unbiased Estimation and Prediction under a Selection Model.

Hmielowski, T., 2018. Assessing the progress of plant breeding programs. CSA News 63, 6–8. https://doi.org/10.2134/csa2018.63.0902

Irani, S., Todd, C.D., 2018. Exogenous allantoin increases Arabidopsis seedlings tolerance to NaCl stress and regulates expression of oxidative stress response genes. J. Plant Physiol. 221, 43–50. https://doi.org/10.1016/j.jplph.2017.11.011

Kaiser, H.F., 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika 23, 187–200. https://doi.org/10.1007/BF02289233

Kataya, A.R.A., Schei, E., Lillo, C., 2016. Towards understanding peroxisomal phosphoregulation in Arabidopsis thaliana. Planta 243, 699–717. https://doi.org/10.1007/s00425-015-2439-5

Kazmi, R.H., Khan, N., Willems, L.A.J., Van Heusden, A.W., Ligterink, W., Hilhorst, H.W.M., 2012. Complex genetics controls natural variation among seed quality phenotypes in a

recombinant inbred population of an interspecific cross between Solanum lycopersicum×Solanum pimpinellifolium. Plant, Cell Environ. 35, 929–951. https://doi.org/10.1111/j.1365-3040.2011.02463.x

Kazmi, R.H., Willems, L.A.J., Joosen, R.V.L., Khan, N., Ligterink, W., Hilhorst, H.W.M., 2017. Metabolomic analysis of tomato seed germination. Metabolomics 13. https://doi.org/10.1007/s11306-017-1284-x

Khan, N., Kazmi, R.H., Willems, L.A.J., van Heusden, A.W., Ligterink, W., Hilhorst, H.W.M., 2012. Exploring the Natural Variation for Seedling Traits and Their Link with Seed Dimensions in Tomato. PLoS One 7, e43991. https://doi.org/10.1371/journal.pone.0043991

Kimura, S., Sinha, N., 2008. Tomato (Solanum lycopersicum): A model fruit-bearing crop. Cold Spring Harb. Protoc. 3. https://doi.org/10.1101/pdb.emo105

Kong, D., Zhu, Y., Wu, H., Cheng, X., Liang, H., Ling, H.Q., 2008. AtTHIC, a gene involved in thiamine biosynthesis in Arabidopsis thaliana. Cell Res. 18, 566–576. https://doi.org/10.1038/cr.2008.35

Kwak, K.J., Kim, B.M., Lee, K., Kang, H., 2017. quatre-quart1 is an indispensable U12 intron-containing gene that plays a crucial role in Arabidopsis development. J. Exp. Bot. 68, 2731–2739. https://doi.org/10.1093/jxb/erx138

Lahmy, S., Guilleminot, J., Schmit, A.-C., Pelletier, G., Chaboute, M.-E., Devic, M., 2007. QQT proteins colocalize with microtubules and are essential for early embryo development in Arabidopsis. Plant J. 50, 615–626. https://doi.org/10.1111/j.1365-313X.2007.03072.x

Lippman, Z.B., Semel, Y., Zamir, D., 2007. An integrated view of quantitative trait variation using tomato interspecific introgression lines. Curr. Opin. Genet. Dev. https://doi.org/10.1016/j.gde.2007.07.007

Liu, L.Y., Li, N., Yao, C.P., Meng, S.S., Song, C.P., 2013. Functional analysis of the ABA-responsive protein family in ABA and stress signal transduction in Arabidopsis. Chinese Sci. Bull. 58, 3721–3730. https://doi.org/10.1007/s11434-013-5941-9

Lu, D., Wang, T., Persson, S., Mueller-Roeber, B., Schippers, J.H.M., 2014. Transcriptional control of ROS homeostasis by KUODA1 regulates cell expansion during leaf development. Nat. Commun. 5. https://doi.org/10.1038/ncomms4767

Miyazaki, Y., Jikumaru, Y., Takase, T., Saitoh, A., Sugitani, A., Kamiya, Y., Kiyosue, T., 2016. Enhancement of hypocotyl elongation by LOV KELCH PROTEIN2 production is mediated by auxin and phytochrome-interacting factors in Arabidopsis thaliana. Plant Cell Rep. 35, 455–467. https://doi.org/10.1007/s00299-015-1896-4

Nakata, M.T., Tameshige, T., Takahara, M., Mitsuda, N., Okada, K., 2018. The functional balance between the WUSCHEL-RELATED HOMEOBOX1 gene and the phytohormone auxin is a key factor for cell proliferation in arabidopsis seedlings. Plant Biotechnol. 35, 141–154. https://doi.org/10.5511/plantbiotechnology.18.0427a

Nic-Can, G.I., López-Torres, A., Barredo-Pool, F., Wrobel, K., Loyola-Vargas, V.M., Rojas-Herrera, R., De-la-Peña, C., 2013. New Insights into Somatic Embryogenesis: leafy cotyledon1, baby boom1 and wuschel-related homeobox4 are epigenetically regulated in Coffea canephora. PLoS One 8. https://doi.org/10.1371/journal.pone.0072160

Nito, K., Kamigaki, A., Kondo, M., Hayashi, M., Nishimura, M., 2007. Functional Classification of Arabidopsis Peroxisome Biogenesis Factors Proposed from Analyses of Knockdown Mutants. Plant Cell Physiol. 48, 763–774. https://doi.org/10.1093/pcp/pcm053

Oliveira, I.C.M., Marçal, T.D.S., Bernardino, K.D.C., Ribeiro, P.C.D.O., Parrella, R.A.D.C., Carneiro, P.C.S., Schaffert, R.E., Carneiro, J.E.D.S., 2019. Combining ability of biomass sorghum lines for agroindustrial characters and multitrait selection of photosensitive hybrids for energy cogeneration. Crop Sci. 59, 1554–1566. https://doi.org/10.2135/cropsci2018.11.0693

Patterson, H.D., Thompson, R., 1971. Recovery of Inter-Block Information when Block Sizes are Unequal. Biometrika 58, 545. https://doi.org/10.2307/2334389

Pessoa, H. P. 2018. Linhagens de introgressão derivadas de Solanum pennellii: tolerância ao déficit hídrico durante a germinação e no estádio de plântula. MSc dissertation, Universidade Federal de Viçosa, Departamento de Fitotecnia.

Przybyla-Toscano, J., Roland, M., Gaymard, F., Couturier, J., Rouhier, N., 2018. Roles and maturation of iron–sulfur proteins in plastids. J. Biol. Inorg. Chem. https://doi.org/10.1007/s00775-018-1532-1

Rocha, J.R. do A.S. de C., Machado, J.C., Carneiro, P.C.S., 2018. Multitrait index based on factor analysis and ideotype-design: proposal and application on elephant grass breeding for bioenergy. GCB Bioenergy 10, 52–60. https://doi.org/10.1111/gcbb.12443

Rocha, J.R. do A.S. de C., Nunes, K.V., Carneiro, A.L.N., Marçal, T. de S., Salvador, F.V., Carneiro, P.C.S., Carneiro, J.E.S., 2019. Selection of superior inbred progenies toward the common bean ideotype. Agron. J. 111, 1181–1189. https://doi.org/10.2134/agronj2018.12.0761

Rosental, L., Perelman, A., Nevo, N., Toubiana, D., Samani, T., Batushansky, A., Sikron, N., Saranga, Y., Fait, A., 2016. Environmental and genetic effects on tomato seed metabolic balance and its association with germination vigor. BMC Genomics 17. https://doi.org/10.1186/s12864-016-3376-9

Schouten, H.J., Tikunov, Y., Verkerke, W., Finkers, R., Bovy, A., Bai, Y., Visser, R.G.F., 2019. Breeding Has Increased the Diversity of Cultivated Tomato in The Netherlands. Front. Plant Sci. 10, 1606. https://doi.org/10.3389/fpls.2019.01606

Shrestha, S., Knapič, M., Žibrat, U., Deleuran, L.C., Gislum, R., 2016. Single seed near-infrared hyperspectral imaging in determining tomato (Solanum lycopersicum L.) seed quality in association with multivariate data analysis. Sensors Actuators, B Chem. 237, 1027–1034. https://doi.org/10.1016/j.snb.2016.08.170

Silva, L.J. da, Medeiros, A.D. de, Oliveira, A.M.S., 2019. SeedCalc, a new automated R software tool for germination and seedling length data processing. J. Seed Sci. 41, 250–257. https://doi.org/10.1590/2317-1545v42n2217267

Silva, M.J., Carneiro, P.C.S., de Souza Carneiro, J.E., Damasceno, C.M.B., Parrella, N.N.L.D., Pastina, M.M., Simeone, M.L.F., Schaffert, R.E., da Costa Parrella, R.A., 2018. Evaluation of the potential of lines and hybrids of biomass sorghum. Ind. Crops Prod. 125, 379–385. https://doi.org/10.1016/j.indcrop.2018.08.022

Stührwohldt, N., Hartmann, J., Dahlke, R.I., Oecking, C., Sauter, M., 2014. The PSI family of nuclear proteins is required for growth in arabidopsis. Plant Mol. Biol. 86, 289–302. https://doi.org/10.1007/s11103-014-0229-2

Tandon, J.P., Jain, H.K., 2004. Plant Ideotype: The Concept and Application, in: Plant Breeding. Springer Netherlands, pp. 585–600. https://doi.org/10.1007/978-94-007-1040-5_25

Toubiana, D., Semel, Y., Tohge, T., Beleggia, R., Cattivelli, L., Rosental, L., Nikoloski, Z., Zamir, D., Fernie, A.R., Fait, A., 2012. Metabolic profiling of a mapping population exposes new insights in the regulation of seed metabolism and seed, fruit, and plant relations. PLoS Genet. 8. https://doi.org/10.1371/journal.pgen.1002612

Wang, H., Yan, M., Xiong, M., Wang, P., Liu, Y., Xin, Q., Wan, L., Yang, G., Hong, D., 2020. Genetic dissection of thousand-seed weight and fine mapping of cqSW.A03-2 via linkage and association analysis in rapeseed (Brassica napus L.). Theor. Appl. Genet. https://doi.org/10.1007/s00122-020-03553-9

Woyann, L.G., Meira, D., Zdziarski, A.D., Matei, G., Milioli, A.S., Rosa, A.C., Madella, L.A., Benin, G., 2019. Multiple-trait selection of soybean for biodiesel production in Brazil. Ind. Crops Prod. 140. https://doi.org/10.1016/j.indcrop.2019.111721

Xi, D.-M., Liu, W.-S., Yang, G.-D., Wu, C.-A., Zheng, C.-C., 2010. Seed-specific overexpression of antioxidant genes in Arabidopsis enhances oxidative stress tolerance during germination and early seedling growth. Plant Biotechnol. J. 8, 796–806. https://doi.org/10.1111/j.1467-7652.2010.00509.x

Yobi, A., Bagaza, C., Batushansky, A., Shrestha, V., Emery, M.L., Holden, S., Turner-Hissong, S., Miller, N.D., Mawhinney, T.P., Angelovici, R., 2020. The complex response of free and bound amino acids to water stress during the seed setting stage in Arabidopsis. Plant J. tpj.14668. https://doi.org/10.1111/tpj.14668

CHAPTER II - Uncovering tomato candidate genes associated with drought tolerance using *Solanum pennellii* introgression lines

ABSTRACT

Tomato plants are sensitive to drought stress throughout their growth cycle. To be considered drought-tolerant, a cultivar should display tolerance at all developmental stages. This study aimed to evaluate whether *Solanum pennellii* introgression lines (ILs) previously selected as drought-tolerant during germination/seedling growth maintained this tolerance in the vegetative/reproductive stage. We then investigated these ILs to uncover candidate genes. The plants were subjected to two different environmental conditions: well-watered and drought-stressed (water withheld for ≤ 20 d after flowering). Phenotyping for morphological, physiological, fruit quality, and yield-related traits was performed, and the data analyzed using a mixed-model approach. The genotypes were then ranked according to their distance from the drought-tolerant ideotype using a multi-trait index based on factor analysis and genotype–ideotype distance (FAI-BLUP index). Next, introgressed segments of interest were identified for candidate gene identification using the tomato IL population map provided by the SOL Genomics Network. Significant genotypic differences were found in the yield, water content, mean weight, length, and width of the fruit, the percentage of fruits displaying blossom-end rot, and titratable acidity. The drought-tolerance ideotype was built considering the maximum values for the fruit water content, number of fruits, mean fruit weight, and yield, minimum values for blossom-end rot, and mean values for titratable acidity. IL 1-4-18, IL 7-4-1, IL 7-1, IL 7-5-5, and IL 1-2 were ranked above M-82 and therefore considered drought-tolerant during the vegetative/reproductive stage. IL 1-4-18 and IL1-2 sustained drought tolerance displayed during germination/seedling growth into the vegetative/reproductive stage. The following candidate genes associated with drought tolerance were identified: *At1g55870, At1g55840, At3g06050, At3g12630, At3g23590, At3g26060, At3g43230, At3g62130, At3g63520, and At4g38810*. Because they are already associated with genetic markers, they can be transferred to elite tomato cultivars through marker-assisted technology after validation.

Keywords: water deprivation, abiotic stress, *Solanum lycopersicum*, breeding.

1. INTRODUCTION

Solanum lycopersicum L., a domesticated tomato, is a popular vegetable crop produced and consumed worldwide, and is one of the world's most economically important vegetables (FAO, 2021). Tomato plants are sensitive to drought stress from seed germination to harvest (Foolad, Zhang, & Subbiah, 2007), demanding high water supply throughout the growing cycle (Ayankojo et al., 2018; Goyal & Sharma, 2018). Even short periods of drought might cause tomato growers qualitative and quantitative losses in production (George et al., 2013; Zhou et al., 2019). Overall, water supply shortages can adversely affect tomato plant growth, physiological characteristics, and yield (George et al., 2013; Yang et al., 2017; Zhou et al., 2019). Therefore, anticipating farmer needs, the development of tomatoes possessing drought tolerance has gained attention in tomato breeding programs (Dariva et al., 2021; Mangat et al., 2021; Zhao et al., 2021).

One of the challenges that delay the delivery of drought-tolerant cultivars is that this trait seems to be a stage-specific phenomenon (Anjum et al., 2017). Nevertheless, an ideal drought-tolerant genotype must display drought resistance at all developmental stages of plant growth (Ebert & Schafleitner, 2015). Therefore, specific stages throughout the ontogeny of the plant must be evaluated separately to assess drought tolerance and to identify its genetic components (Foolad, 2007).

The green-fruited wild relatives *Solanum chilense* L. and *Solanum pennellii* Corr. are the main sources of genetic variation in tomato drought tolerance (Bolger et al., 2014; Foolad, 2007; Moyle & Muir, 2010). However, using *S. chilense* in breeding programs is not viable because it displays several barriers to crossing with cultivated tomatoes (Kumar et al., 2019). In contrast, *S. pennellii* is a crossable wild relative, but direct crosses with cultivated tomatoes can lead to the linkage drag of undesirable traits related to yield and fruit quality (Ebert & Schafleitner, 2015).

A viable way to overcome this problem is through the exploitation of the *S. pennellii* genome available with a background of domesticated tomatoes in the collection of introgression lines (ILs) developed by Eshed and Zamir (1995). In this collection, defined genomic segments of *S. pennellii* have replaced homologous regions in the commercial cultivar M-82, providing complete coverage of the wild species genome. ILs are a valuable genetic resource for studying drought tolerance and identifying the genomic regions associated with it, because any phenotypic differences found between an IL and the recurrent parent M-82 when subjected to

drought conditions can be solely attributed to the introgressed chromosomal segment of the drought-tolerant *S. pennellii* (Lippman et al., 2007; Schauer et al., 2006).

In this study, to aid and accelerate the development of drought-tolerant cultivars, we evaluated *S. pennellii* ILs that were previously selected based on their level of drought tolerance during their germination and seedling stages (Pessoa, 2018). The genotypes were exposed to drought by withholding water during the vegetative/reproductive stage and were ranked according to their drought tolerance. Our main goal was to provide information on whether the previously selected drought-resistant genotypes can sustain their tolerance when drought stress is applied in the vegetative/reproductive stage and should, therefore, be incorporated in the next phases of tomato breeding programs. Moreover, the introgressed *S. pennellii* fragments of the genotypes selected as drought-tolerant will be analyzed to identify candidate genes associated with this trait in the vegetative and reproductive stages.

MATERIALS AND METHODS

2.1. PLANT MATERIALS AND GREENHOUSE EXPERIMENTS

Two experiments with similar conditions, except for irrigation, were conducted simultaneously to evaluate the response of genotypes to drought stress. In the first experiment, hereafter "environmental condition 1," plants were kept in non-stressed conditions throughout the growing season, with the soil water content at field capacity. In the second experiment, hereafter "environmental condition 2," plants were challenged with drought by withholding water for up to 20 d after flowering (Ahammed et al., 2020).

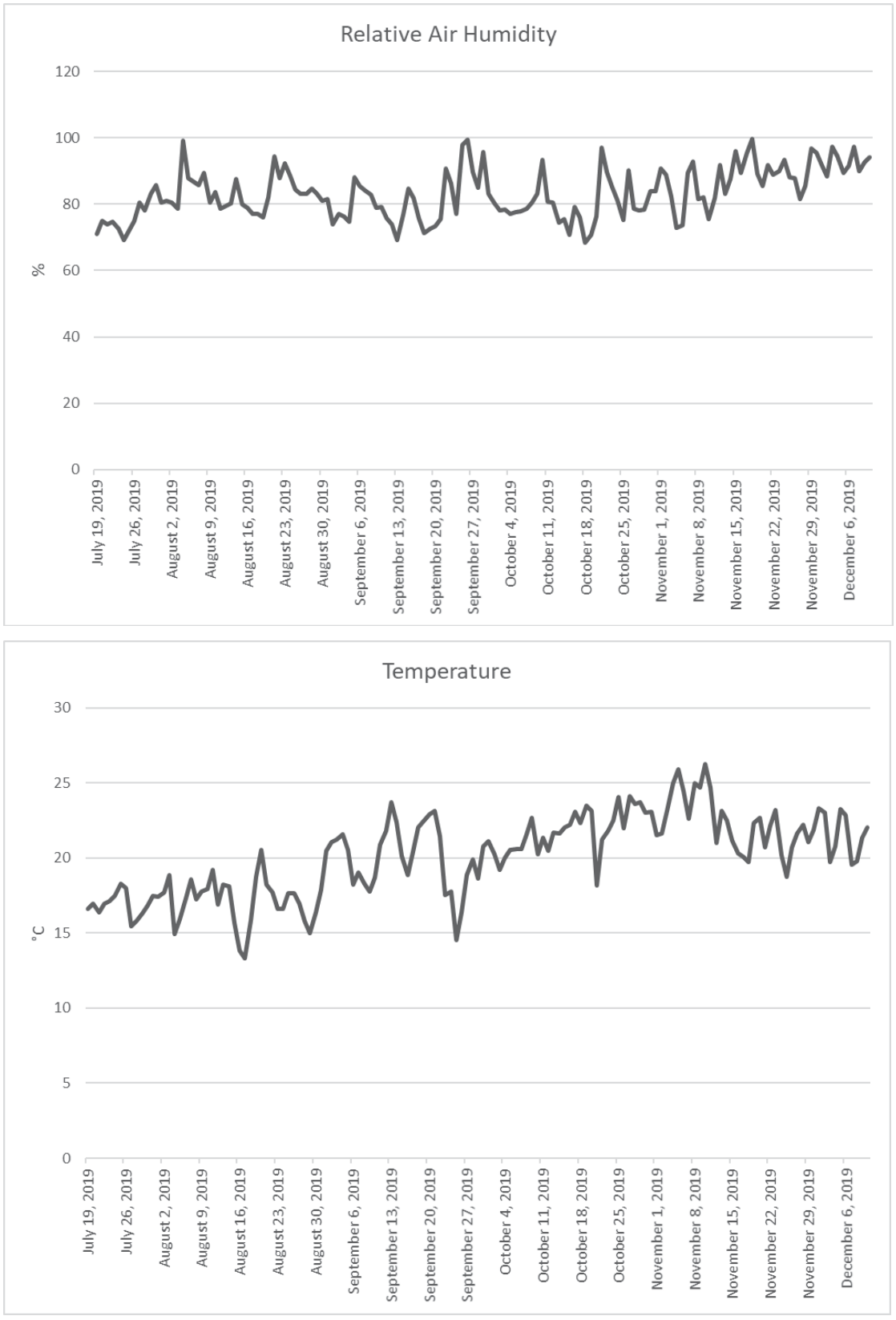
The plant material was chosen based on a previous study (Pessoa et al., 2018) that evaluated the drought tolerance of ILs from the collection developed by Eshed and Zamir (1995) during the germination and seedling stages. The five ILs previously ranked as the most drought-tolerant (IL 1-4-18, IL 2-3, IL 1-2, IL 9-2, and IL 10-1) and the five ranked as the most drought-sensitive (IL 8-3, IL 7-4-1, IL 7-5-5, IL 9-3, and IL 7-1) together with one of their parents, the commercial variety M-82, were evaluated in this study (Table 1).

Table 1. List of genotypes evaluated

Drought tolerance at germination and seedling stages*	**Genotype**	**Introgression line**
Drought tolerant	1	IL 1-4-18
	2	IL 2-3
	3	IL 1-2
	4	IL 9-2
	5	IL 10-1
Drought sensitive	6	IL 8-3
	7	IL 7-4-1
	8	IL 7-5-5
	9	IL 9-3
	10	IL 7-1
Parent	**11**	**cv. M-82**

*Drought tolerance and/or sensitivity was determined in a previous study (Pessoa, 2018).

The greenhouse experiment was conducted in the Research and Extension Farm Unit *Horta Velha* at the Universidade Federal de Viçosa, Viçosa, MG, Brazil (20° 45′ 14″ S; 42° 52′ 53″ W; 648.74 m altitude) from July to December 2019. Temperature (T, °C) and relative air humidity (RH, %) inside the greenhouse were monitored daily using a meteorological station, E4000 (Irriplus ® Scientific Equipment) - Supplementary Figure 1).



Supplementary figure 1. Relative humidity, in percentage, and temperature in °C throughout the experiment.

Seeds were sown in polystyrene trays of 128 cells, each containing Tropstrato substrate (Vida Verde, BRA). Plants at the three to four true leaf stage were transplanted into 15-L pots (one plant per pot) containing a mixture of soil, sand, and composted cow manure (3:1:1). The soil texture was classified as sandy clay (sand, 52.9 %; silt, 4.5 %; clay, 42.6 %). Fertilization and fungicide/insecticide applications were based on crop recommendations. The plants were tied to bamboo sticks placed inside each pot to prevent falling.

Each experiment was arranged in a randomized block design, with three replicates. Each replicate consisted of the mean value of three plants arranged side-by-side.

2.2 IRRIGATION MANAGEMENT AND CALCULATION

In accordance with the proposal by Wudiri & Henderson (1985) that tomato plants are most sensitive to drought during flowering and fruit setting, recently confirmed by Cui et al. (2019), this was the stage when drought conditions were imposed on the plants in environmental condition 2.

Under environmental condition 1, the soil water content was maintained at 100 % available soil water (ASW) during the entire growing season. Under environmental condition 2, after the flowering stage, plants were subjected to drought stress by withholding water until the soil matric potential approached the wilting point ~ 0 % ASW (~ −1500 kPa).

Irrigation management and calculations followed the methodology described by Dariva et al. (2020), as detailed bellow. A soil sample was taken to estimate the soil water retention curve parameters based on the Van Genuchten (1980) equation using SWRC Fit software (Seki, 2007). To estimate the water tension value corresponding to 100 % ASW, we assumed that field capacity was reached at a soil matric potential of −1,500 kPa (Bernardo et al., 2019). To estimate when to stop water deprivation under environmental condition 2, we assumed that the wilting point would be reached at a soil matric potential of −1500 kPa (Bernardo et al., 2019). Therefore, the plants were re-watered when the soil matric potential was estimated to be approximately −1490 kPa. After the drought stress period, the plants under both environmental conditions were kept well-watered until the end of the growing season.

Before filling them with soil, all pots were weighed to determine their recipient weight (W_{rec}). In addition, the bamboo stick used in each pot was weighed to determine the tutor's weight (W_{tutor}). Subsequently, all pots were filled with the same dry soil weight (W_{ds}). The water

weight (W_{water}) of the first irrigation was determined by multiplying W_{ds} by the soil water content (kg/kg) at a soil water potential of −33 kPa.

The soil water content was monitored daily by weighing each pot. For the daily irrigations, the applied W_{water} was determined as total pot weight ($W_{totalpot}$) minus the pot weight measured on that day. The total pot weight was calculated as follows: $W_{totalpot} = W_{rec} + W_{ds} + W_{water} + W_{plant} + W_{tutor}$, where W_{rec} is the recipient weight, W_{ds} is the dry soil weight, W_{water} is the water weight, W_{plant} is the plant weight, and M_{tutor} is the tutor weight. W_{plant} was determined by weighing same-aged spare plants grown during the experiment. One spare plant was harvested and weighed every 10 d to adjust the W_{plant} value.

2.3. PHENOTYPING

2.3.1 PHYSIOLOGICAL AND MORPHOLOGICAL TRAITS

Leaf water potential (LWP) was measured predawn (3:00–5:00) using a pressure chamber (model 3000; Soil Moisture, Santa Bárbara, CA, USA), according to the method described in Scholander et al. (1965). The LWP value for each repetition was the mean value of two leaflets.

The leaf relative water content (LWC) was determined at the end of the drought stress period for three leaf samples randomly picked from the middle portion of each plant. LWC was calculated according to Egea et al. (2018) using the following formula:

$$LWC = (FW - DW)/(SW - DW)$$

where FW is the fresh weight, DW is the dry weight obtained after oven-drying for 48 h at 80 °C, and SW is the saturation weight determined after 24 h of re-saturation in tap water.

Shoot dry matter was determined for the whole shoot (leaves and stems) at the end of the experiment by oven-drying for 48 h at 80 °C. Shoot water content (SWC) was determined for the whole shoot (leaves and stems) at the end of the experiment using the following formula (Egea et al., 2018):

$$SWC = (FW - DW)/(SW - DW)$$

where FW is the fresh weight and DW is the dry weight obtained after oven-drying for 48 h at 80 °C.

Fruit water content (FWC) was determined for a sample of three to four fully ripened fruits per plant picked randomly at the end of the experiment, according to the following formula (Egea et al., 2018):

$$FWC = (FW - DW)/(SW - DW)$$

where FW is the fresh weight, and DW is the dry weight obtained after oven-drying for 48 h at 80 °C.

2.3.2 YIELD

Tomato fruits were harvested when fully ripe during the course of the experiment. The fresh weight of all fruits from each plant was measured. The yield parameters consisted of total plant yield (Kg plant^{-1}), yield, mean fruit weight (MFW, g), and the number of fruits per plant (N_fruits).

2.3.3 FRUIT QUALITY ATTRIBUTES

The visual quality attributes—fruit length (FL) and width (FW)—of all harvested fruits were measured using a digital pachymeter. Each fruit was visually evaluated to assess the presence of blossom endrot (BER), scars (SC), and cracking (CR), and the results were expressed in terms of the percentage of fruits with defects.

Organoleptic quality measurements were performed randomly on four fully ripened fruits per plant. All four fruits per plant were homogeneous in size and color and were harvested on the same day; fruit quality assessment began immediately after harvest.

Fruit firmness, described as the mean maximum penetration force required for pericarp rupture and expressed in Newtons (N), was measured using a digital penetrometer (model PDF-200, Soilcontrol, USA) fitted with a cylindrical stainless-steel probe with a round tip (Ø 8 mm). Two measurements, located 180º apart from one another, were taken in the equatorial region of each fruit.

After firmness measurement, the four selected fruits were macerated together in a blender to produce tomato juice, which was used to determine the total acidity (pH), total soluble solids (TSS), and titratable acidity (TA). The pH values of the juice samples were determined immediately using a benchtop pH meter (model pH 21, Hanna Instruments, Italy). TSS, expressed as °Brix, was determined using a digital refractometer (model HI 96801, Hanna

Instruments, Italy). For TA measurements, samples of approximately 5 g of tomato juice were transferred to 100-mL volumetric flasks, which were then filled to capacity with distilled water. An aliquot of 10 mL of this solution was then titrated with an NaOH solution (0.005 mol L^{-1}) using 1 % phenolphthalein as an indicator. TA values were expressed as % citric acid and were obtained using the following formula:

$$TA = [(mL\ NaOH) \times (0.005\ N\ NaOH) \times (0.064\ mEq\ acid\ citric\ factor) \times (correction - factor) \times 100] / grams\ of\ juice\ sample$$

Finally, the TSS/TA ratio was calculated as an indicator of flavor, as described by Kader et al. (1978).

2.4. STATISTICAL ANALYSIS

A mixed-model methodology was adopted for statistical analyses via REML/BLUP (restricted residual maximum likelihood/best linear unbiased prediction), according to Patterson and Thompson (1971) and Henderson (1975), using R software and the Sommer package.

The statistical model was denoted by:

$$y = X\tau + Z_g\mu_g + Z_i\mu_i + e \quad (2)$$

where Y is the phenotypic data vector; $\tau$is the fixed effects vector (overall mean, environment, and blocks); $\mu_g$is the genotype effects vector (random), with $\mu_g \sim N(0, \sigma_g^2)$; μ_i is the genotype × experiment interaction (random), with $\mu_i \sim N(0, \sigma_i^2)$; e is the residual effects vector (random), with $e \sim N(0, R)$; and X, Zg, and Zi are incidence matrices for τ, μ_g, and μ_i, respectively. R is a diagonal matrix with residual variances, that is, $R = \bigotimes_{i=1}^{t} \sigma_{e_i}^2$ where $\sigma_{e_i}^2$ is the residual variance in experiment i and t is the number of environmental conditions (in this case = 2).

For the random effects of the model, the significance of the likelihood ratio test was assessed using the chi-square statistic with one degree of freedom. Genetic values (BLUP means) were predicted for each of the 11 genotypes based on the 37 evaluated traits and two environmental conditions. For traits where significant genotype effects were detected, the BLUP means were used to estimate Pearson's linear correlation coefficients.

2.4.1 SINGLE-TRAIT AND MULTI-TRAIT IL RANKINGS

The genetic values (BLUP means) of traits that showed significant genotype effects ($P < 0.05$) were submitted to single-trait ranking, according to the designed ideotype: maximum values for FWC, N_Fruit, MFW, and yield; minimum values for BER; and mean values for TA. For each trait, the percentage difference of each IL was calculated relative to the commercial cultivar M-82. Subsequently, a multi-trait index based on factor analysis and genotype–ideotype design, known as the FAI-BLUP index, was used to rank the genotypes. Principal component analysis, factor analysis, ideotype determination, and genotype–ideotype distance calculation were performed in R software using the FAI-BLUP index routine developed by Rocha et al. (2018).

Principal component analysis was used to extract the genetic correlation matrix factorial loads obtained by the genetic values. The varimax criterion described by Kaiser (1958) was used for analytic rotation, and the weighted least squares method described by Bartlett (1938) was used to calculate the factor scores. The number of ideotypes was defined based on a combination of desirable and undesirable factors for the selection objective. The following algorithm provides the number of ideotypes:

$$NI = 2^n$$

in which, NI is the number of ideotypes and n is the number of factors.

After ideotype determination, genotype–ideotype distances were estimated and converted into spatial probabilities, enabling genotype ranking. The following algorithm was used.

$$P_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{i=1;j=1}^{i=n;j=m} \frac{1}{d_{ij}}}$$

where P_{ij} is the probability that the i^{th} genotype (i = 1, 2,..., n) is similar to the j^{th} ideotype (j = 1, 2,..., m) and d_{ij} is the genotype–ideotype distance from the i^{th} genotype to the j^{th} ideotype, based on the standardized mean Euclidean distance.

2.5. CANDIDATE GENE IDENTIFICATION

The genomic segments of *S. pennellii* introgressed into the ILs selected by the FAI-BLUP index were further analyzed to identify candidate genes for drought tolerance, adapting

the methodology described previously (Toubiana et al., 2012). We used a map of the tomato IL population provided by the SOL Genomics Network (http://solgenomics.net/). This map displays individual chromosomes with restriction sites for the different introgressed segments, as well as all identified marker genes.

We adapted the script developed in Toubiana et al. (2012) to automatically identify all the marker genes from each selected IL in the HTML code. The functionality of each identified gene was inferred using information on their respective orthologs in the *Arabidopsis thaliana* genome from the TAIR database (https://www.arabidopsis.org/). The identification of candidate genes for drought tolerance was based on the relevance of their functionality to this process.

3. RESULTS

The effect of environment was significant for LWP (p-value > 0.05), meaning that the mean LWPs of the two environmental conditions were different. Figure 1 shows a boxplot of LWP under environmental conditions 1 and 2. As expected, the values observed for environmental condition 1, in which the plants were well-watered throughout the growing cycle, were higher than those observed for environmental condition 2, in which the plants were subjected to drought stress by water deprivation.

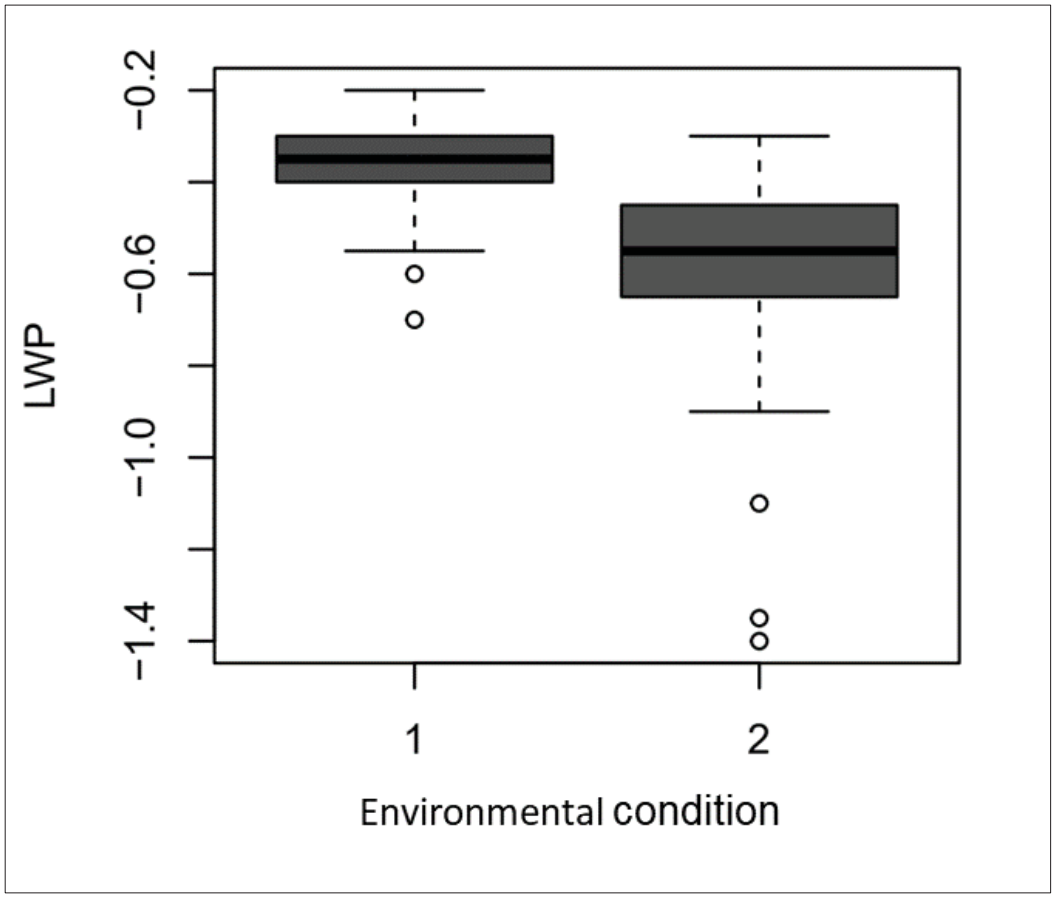


Figure 1. Boxplot of leaf water potential (LWP) in the two environmental conditions, Environmental condition 1 = non-stressed plants, environmental condition 2 = plants sujected to drought stress.

Among the 18 evaluated traits, significant genotypic differences were found for five yield and fruit quality attributes: FWC, MFW, FL, FW, BER, and TA. The effect of the interaction between genotype and environmental conditions was not significant for any of the evaluated traits.

3.1 SPEARMAN'S CORRELATION

Our study on the associations between the yield and fruit quality attributes revealed 19 significant linear correlations, of which 13 were positive and six negative ($p < 0.05$, t-test) (Figure 2). Overall, the yield attributes were positively correlated. As MFW displayed a strong positive correlation with FW ($r = 0.98$), we decided to retain only one of them for further analysis and removed FW. The other correlations among the traits were only moderate ($r < 0.7$); hence, all the other traits were retained. Yield showed a moderate positive correlation with N_ fruit, FWC ($r = 0.78$), and MFW ($r = 0.75$).

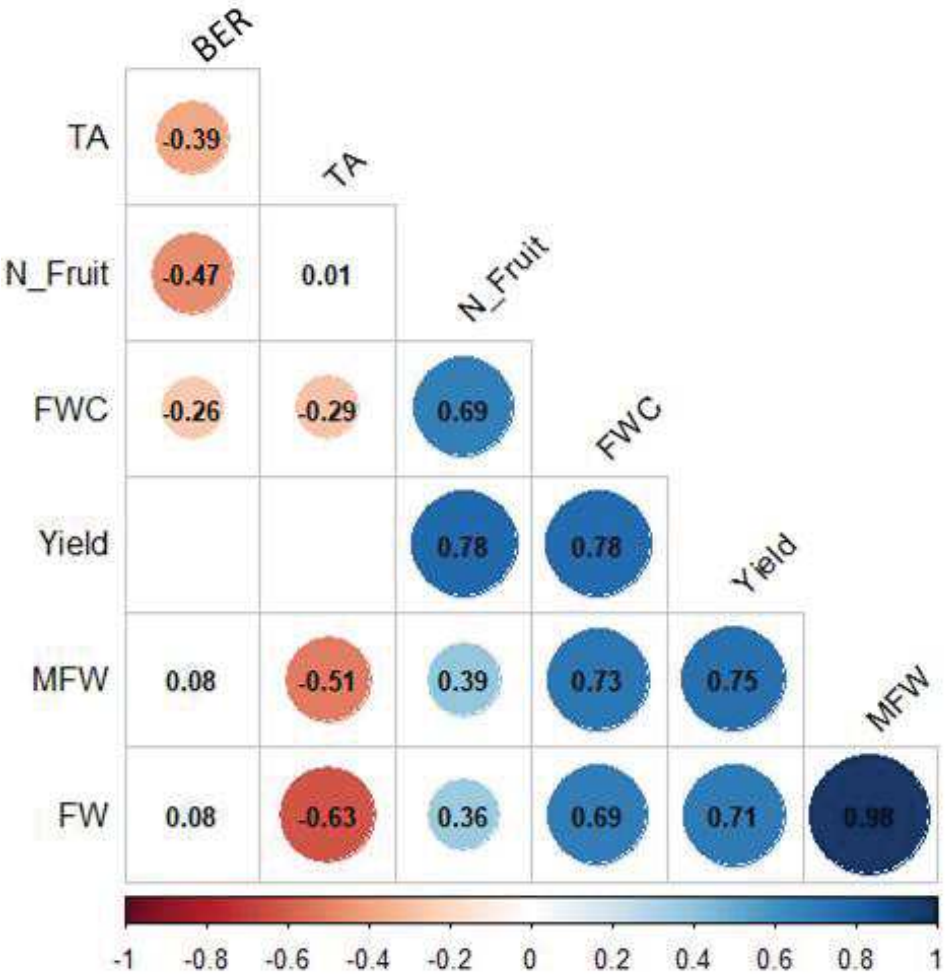


Figure 2. Pearson's correlation matrix of the BLUP means of all traits showing significant genotypic effects. Correlations were validated by the t-test at a 0.05 significance level, and blank squares mean that the correlation was not significant ($p > 0.05$). BER = blossom-end rot, FW = fruit width, FWC = fruit water content, MFW = mean fruit weight, N_fruits = mean number of fruits per plant, TA = titratable acidity, Yield = total yield.

3.2 SIGLE-TRAIT AND MULTI-TRAIT RANKING OF ILS SUBJECTED TO DROUGHT STRESS

Figure 3 displays the genotype ranking according to the ideotype for each trait. The ideotype was designed considering maximum values for FWC, N_Fruit, MFW, and yield, minimum values for BER, and mean values for TA. The chromosomal fragments of *S. pennellii* introgressed into cv. M-82 promoted changes in all the ILs for all traits. IL7-1 stood out in terms of the yield parameters. Compared with its genitor, there was an impressive increase in mean fruit weight (68.83 %) and total yield (32.24 %). This IL also resulted in the highest increase in fruit number. IL 7-4-1 was among the top five genotypes for all traits, and its only trait that was not improved by the introgression of *S. pennellii* was BER. When analyzing all the traits together, it is evident that an introgressed region can increase some traits while decreasing others. In addition, a genotype ranked first for one trait could be ranked last for others. For example, the introgression of *S. pennellii* promoted an increase of 49.14 % in the MFW of IL 1-4-18; this IL was also ranked second for FW and penultimate for BER.

Figure 4 presents the ranking of the 11 evaluated genotypes according to the multi-trait index FAI-BLUP, and the probability of good drought tolerance in relation to the distance to the ideotype. This method ranked the genotypes according to the proposed ideotype (maximum values for FWC, N_Fruit, MFW, and yield; minimum values for BER; and mean values for TA), considering all the parameters simultaneously. The cultivar M-82 was ranked the sixth closest to the drought-tolerant ideotype. Therefore, the five ILs ranked above it were selected as drought-tolerant and investigated for the presence of candidate genes related to this trait. Among the five ILs selected as drought-tolerant, two were previously determined to be drought-tolerant in the germination and seedling stages, IL 1-4-18 and IL-1-2. In addition, the most drought-tolerant plants at the germination and seedling stages (IL 1-4-18) were also the most drought-tolerant during the vegetative/reproductive stage.

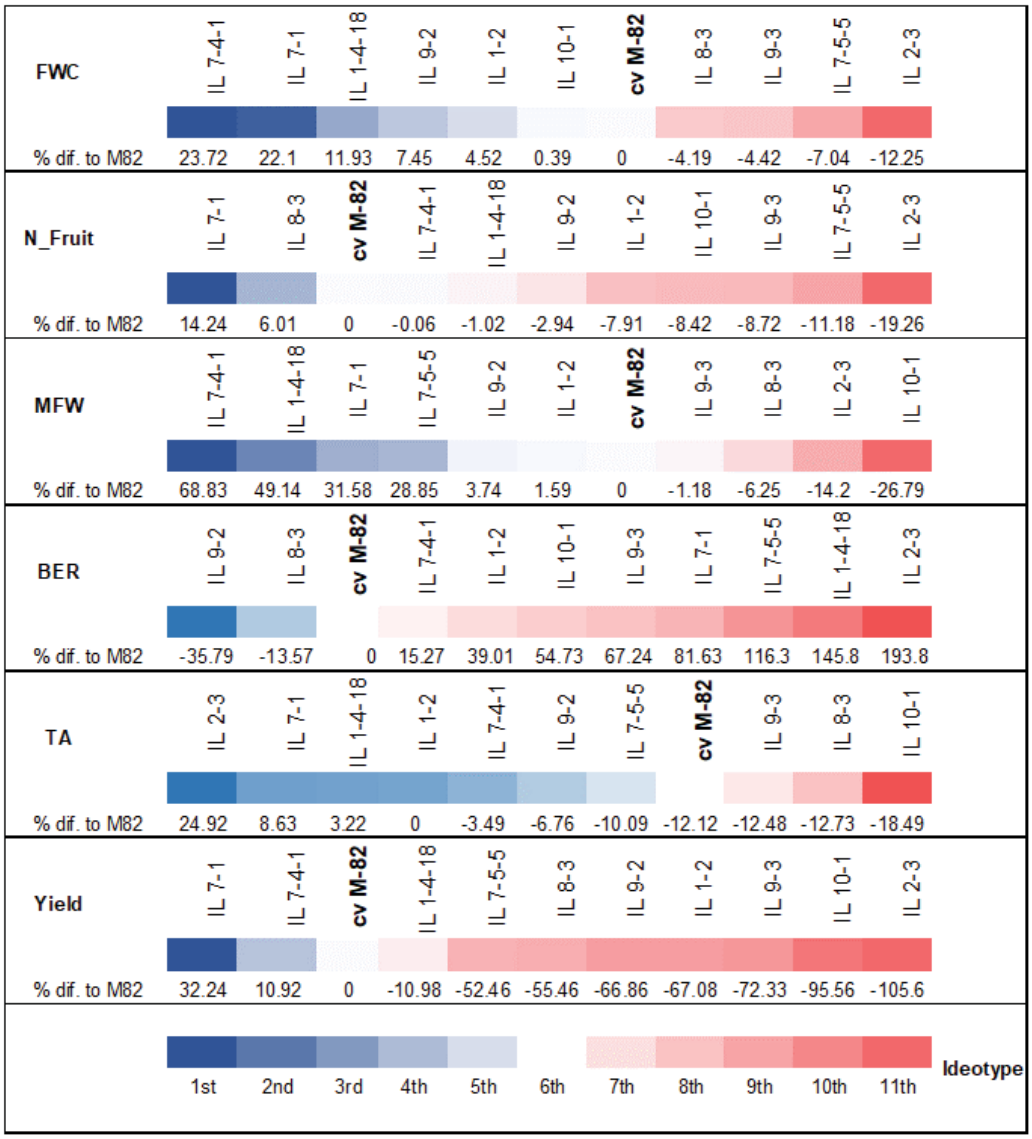


Figure 3. Ranking of genotypes according to the ideotype for each trait and the percentage difference (increase or decrease) in the trait in relation to cv. M-82. The genotypes closest to the ideotype are on the left. The colored gradient indicates the intensity of difference in relation to M-82: the deeper the blue and red colors, the more substantial the changes promoted by the S. pennnelli genome away from the ideotype. BER = blossom-end rot, FWC = fruit water content, MFW = mean fruit weight, N_fruits = mean number of fruits per plant, TA = titratable acidity, Yield = total yield.

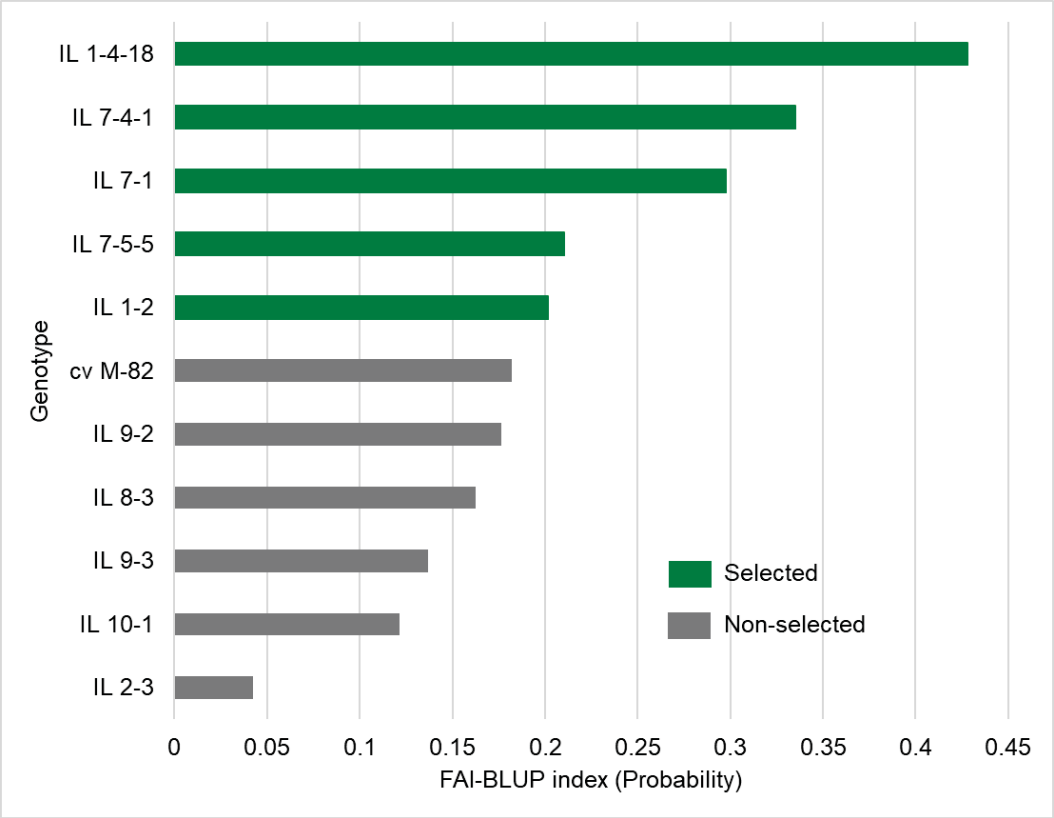


Figure 4. Genotype ranking and selected genotypes using the FAI-BLUP index.

3.3 CANDIDATE GENES ASSOCIATED WITH DROUGHT TOLERANCE

All the genotypes ranked above M-82 (IL 1-4-18, IL 7-4-1, IL 7-1, IL 7-5-5, and IL 1-2) were further investigated for candidate genes associated with drought tolerance. Gene markers associated with the introgressed segments of the selected ILs were identified using the SOL Genomics Network. Their functionality was inferred using information on their respective orthologs in the *A. thaliana* genome. Figure 5 provides a summary of all the candidate genes discovered.

In IL 1-4-18, the gene *At4g38810* was identified. It encodes SnRK2, a protein involved in plant response to water deprivation (Fujita et al., 2009). In IL 7-4-1, *At3g26060* and *At3g43230* were identified. *At3g26060* encodes peroxiredoxin Q, which decomposes peroxides, and is therefore involved in the cellular response to oxidative stress (Jing et al., 2006). *At3g43230* encodes a RING-type zinc-finger family protein involved in the response to ABA (Zang et al., 2015).

In IL 7-1, the genes *At1g55840* and *At1g55870* were identified. *At1g55840* encodes a Sec14p-like phosphatidylinositol transfer family protein that is involved in the defense response

to abiotic stress (Huang et al., 2016). *At1g55870* encodes a poly(A)-specific ribonuclease AtPARN. The expression of AtPARN is upregulated by ABA or stress treatment, acting upstream of or within the response to ABA and osmotic stress (Zhang et al., 2008). No genes related to drought tolerance were found in the selected IL 7-5-5.

In IL 1-2, the genes *At3g06050, At3g12630, At3g23590, At3g62130,* and *At3g63520* were identified as candidates. *At3g06050* encodes ATPRXIIF, which is involved in oxidative stress response (Geigenberger et al., 2017). *At3g12630* encodes stress-associated protein 5 (SAP5), a protein that positively regulates stress responses in Arabidopsis and responds to water deprivation (Kang et al., 2011). *At3g23590* encodes REF4-RELATED 1, which regulates the phenylpropanoid metabolic process related to plant response to abiotic stress (Sharma et al., 2019). *At3g62130* encodes L-cysteine desulfhydrase (LCD), which is involved in plant response to drought (Jin et al., 2013). Finally, *At3g63520* encodes 9-cis-epoxycarotenoid dioxygenase, which responds to water deprivation (Iuchi et al., 2001).

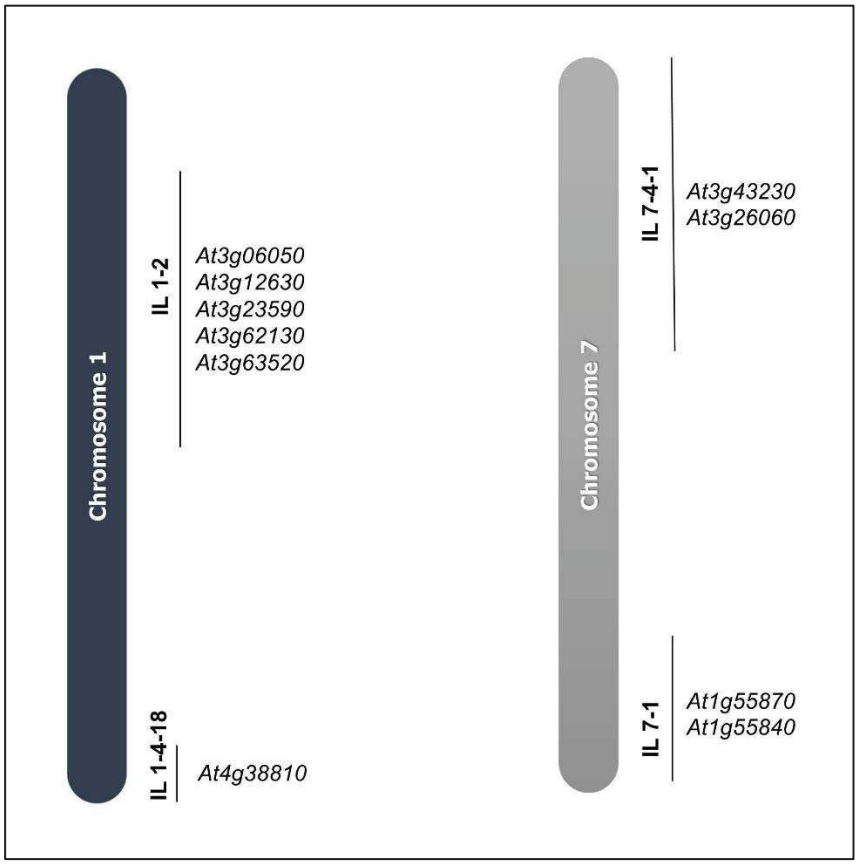


Figure 5. Candidate genes associated with drought tolerance in the selected ILs.

4. DISCUSSION

An ideal drought-tolerant genotype must display drought tolerance at all developmental stages of plant growth (Ebert & Schafleitner, 2015). Therefore, specific stages throughout the ontogeny of the plant must be evaluated separately (Foolad, 2007). Here, we selected genotypes previously screened and ranked for drought tolerance during seed germination and seedling growth and evaluated them during the vegetative and reproductive stages under two water regimes (100 % ASW and drought exposure effected by withholding water for up to 20 d).

The significant difference among the environments observed for LWP ($p < 0,05$) highlights that the adopted methodology was efficient in promoting drought stress in the plants from which water was withheld. Despite the wide range of methodologies proposed to induce drought stress in tomatoes (Foolad, 2007; Dariva et al., 2020; Makalesi, 2020; Vanisree et al., 2021), multiple studies, from the pioneering work of Wudiri & Henderson (1985) to the recent experiments of Cui et al. (2019), suggest that the flowering and fruit-setting stage is when tomato plants are the most sensitive to drought. Therefore, promoting severe water deprivation (soil almost reaches wilting point) during this stage, as we did in our study, proved to be an accurate method of screening and selecting plants exhibiting drought tolerance.

As expected, the plants from environmental condition 2, which were exposed to drought, showed lower LWP than those from environmental condition 1. Predawn LWP usually corresponds to a daily maximum or base water potential, which is presumed to correspond to the equilibrium between the soil and plant water potentials (Sellin, 1996). It is believed that predawn plant water potential is in equilibrium with the “wettest” soil accessed by the roots (Richter, 1997; Ritchie & Hinckley, 1975). Therefore, the predawn LWP was used as a surrogate for the soil water potential. After irrigation suspension in environmental condition 2, evapotranspiration consumed the water available in the soil over the entire period of drought stress. By the end of the period, the soil reached a much lower water potential than that in the pots that were watered daily.

We did not identify significant genotype × environment relationships for any of the evaluated traits. Like most herbaceous plants, 90 % of the total fresh weight of tomato plants is water (Mitchell et al., 1991). For tomato fruits, this percentage is even higher; water accounts for 93–95 % of tomato composition (Foolad &Foolad, 2004). All chemical and biochemical cellular processes, including photosynthesis, take place within an aqueous media. Water deficit immediately stops cell division and enlargement and stem and leaf elongation, processes that

require a minimum turgor pressure. A lack of water also inhibits photosynthesis, closes stomata, and decreases respiration rates and other enzyme-mediated processes (Zsögön et al., 2017). Thus, it was expected that the genotypes under drought stress would show different responses for the evaluated traits than those under watered conditions. The LWP values confirmed that the plants in environmental condition 2 were under drought stress conditions. However, the drought period was short, considering the extent of the growth cycle, which was approximately five months. A hypothesis for the non-interaction between genotype and environment is that after drought stress suspension, the plants subjected to water deficit adjusted their physiological and biochemical machinery to adapt to these conditions. This suggests a valuable set of genetic resources for the further investigation of drought tolerance.

Genotypic differences ($p < 0,05$) were identified for only seven of the 18 evaluated traits. This can be explained by the fact that the 11 genotypes used as plant material were highly genetically related (10 ILs from the collection developed by Eshed & Zamir (1995) and their genitor, the commercial cultivar M-82). Therefore, it is likely that the genomic portions shared among the genotypes, rather than the introgressed fragments, are associated with the traits for which significant genotypic differences were not observed. In this study, the differences promoted by the introgressed fragments are the ones of interest; therefore, only the traits displaying genotypic differences were further analyzed.

The chromosomal fragments of *S. pennellii* introgressed into cv. M-82 promoted changes in fruit quality- and yield-related traits (FWC, MFW, FL, FW, yield, BER, and TA) in the ILs. M-82 was not the first-ranking genotype for any of these, indicating that the genomic introgressions were favorable for some of the ILs. Although *S. pennellii* bears small green fruits, unlike cultivated genotypes, some wild alleles might have opposite effects to those expected when analyzing the wild relative phenotype of a crop (Ebert & Schafleitner, 2015). For example, while examining the function of wild tomato alleles in a background of tomato cultivars through QTL analysis, Eshed and Zamir (1995) identified two ILs with larger fruit sizes than the cultivated parent, and QTL mapping in the IL population revealed that the allele leading to larger fruit size was, surprisingly, contributed by *S. pennellii*. Moreover, pyramiding different genetic loci derived from *S. pennellii* into *S. lycopersicum* lines in a heterozygote state increased fruit yield by 30–50 % under field conditions compared to an elite variety (Gur et al., 2011).

From the single-trait ranking, it is evident that an introgressed region can increase some traits while decreasing others. In addition, a genotype ranked first for one trait could be ranked

last for others. Therefore, single-trait ranking is not helpful in selecting the genotypes best suited to drought conditions during the vegetative and reproductive stages, justifying the use of a multi-trait approach.

The FAI-BLUP index ranked five ILs above M-82, confirming that the *S. pennellii* genomic fragments in these ILs increased their drought tolerance. *S. pennellii* possesses several adaptive mechanisms that ensure its survival in arid environments, including morphophysiological and anatomical modifications in the aerial part of the plant, such as a cuticular composition associated with increased resistance to water flux, a smaller leaf surface area, and greater leaf thickness. Therefore, it was expected that the introgressed *S. pennellii* fragments in the ILs would carry genes that enhance drought tolerance. Moreover, IL 1-4-18 and IL-1-2, which were ranked as drought-tolerant in this study, were also ranked as drought-tolerant in their germination and seedling stages. Drought tolerance is a stage-specific phenomenon that is controlled by many genes with different effects (Egea et al., 2018; Florido Bacallao & Bao Fundora, 2014; Foolad & Foolad, 2004). Therefore, most probably, the genes controlling the drought tolerance of IL 1-4-18 and IL-1-2 during germination and seedling growth were not the same as those controlling this trait in the stages evaluated in our study. Hence, the fragments of *S. pennellii* inherited by these ILs are a rich source of drought tolerance-related genes, with genes acting during seed germination and seedling growth, and genes expressed during the vegetative and reproductive stages. Our results suggest that IL 1-4-18 and IL-1-2 are promising genotypes for inclusion as genitors in breeding programs aimed at developing drought-tolerant tomato cultivars.

IL 1-4-18 was ranked the most drought-tolerant genotype and the gene *At4g38810*, which encodes SnRK2, was identified in this IL. Kulik et al. (2011) described SnRK2 as a crucial regulator of plant responses to abiotic stresses. This protein acts as a positive central regulator of ABA signaling during water stress at the vegetative and reproductive stages (Fujita et al., 2009). Water deficit stress, such as drought, triggers various biochemical and physiological responses in plants, including alterations in gene expression and the accumulation of the phytohormone ABA (Pashkovskiy et al., 2019). ABA regulates diverse plant processes, including the adaptation of plants to water stress (Daszkowska-Golec, 2016; Lim et al., 2015). Numerous drought stress-responsive genes have been reported, many of which are induced by ABA (López-Galiano et al., 2019). SnRK2s pathways regulate the plant response to ABA by the direct phosphorylation of various downstream targets, such as SLAC1, KAT1, AtRbohF, and transcription factors required to express numerous stress response genes (Kulik et al.,

2011). Therefore, the gene discovered in our study might have played a central role in the response of IL 1-4-18 to water deprivation, resulting in drought tolerance.

In IL 7-4-1, the genes *At3g26060* and *At3g43230* were identified. *At3g43230* encodes a RING/FYVE/PHD-type protein. This group of proteins are nuclear ubiquitin E3 ligases. Although they do not seem to be directly involved in transcriptional regulation, they can interact with some transcription factors to regulate their transcriptional activities, resulting in improved abiotic and biotic stress tolerance in several species (Zang et al., 2015). *At3g26060* encodes peroxiredoxin Q, which plays a role in the cellular response to oxidative stress caused by peroxide decomposition. Peroxiredoxins are part of the enzymatic ROS-scavenging system developed by plants to scavenge ROS generated during biotic stress events such as drought (Liebthal et al., 2018). Peroxiredoxin Q, one of the four plant subtypes, is a chloroplast gene associated with responses to several abiotic stresses (Jing et al., 2006). By scavenging peroxides generated under stress, peroxiredoxin Q protects the photosynthetic machinery, enabling photosynthesis to continue in non-optimal conditions. IL 7-4-1 was ranked as the second-most drought-tolerant genotype and showed outstanding behavior for all evaluated traits under drought conditions. *At3g26060* and *At3g43230* may be key genes in the drought-tolerance strategy of this genotype.

In IL 7-1, the genes *At1g55840* and *At1g55870* were identified. *At1g55840* encodes a Sec14p-like phosphatidylinositol transfer family protein. In soybeans, this protein is a component of a stress response pathway that serves to protect adult plants under osmotic stress (Monks et al., 2001). Proteins with a Sec14p-like domain play a role in the lipid regulation of the Rho-mediated signaling pathway, which is involved in water stress tolerance in wild barley (Suprunova et al., 2007). *At1g55870* encodes AtPARN. Although the specific roles of this protein are not well-described, they were found to be expressed when plants were subjected to drought stress (Zhang et al., 2008), and according to Covarrubias & Reyes (2010), they are involved in the post-transcriptional gene regulation of salinity and drought responses. These two genes are good candidates for explaining the drought tolerance of IL 7-1.

IL 7-5-5 was selected as the fourth-closest genotype to the drought tolerance ideotype (Figure 3); however, we did not identify any candidate genes that could justify this result. In our approach, genes with unknown functionalities or those that did not have an *A. thaliana* ortholog were not included in the list of potential candidate genes. Hence, it is possible that one or several of these excluded genes are related to the drought tolerance displayed by IL 7-5-5 in this study.

In IL 1-2, the genes *At3g06050, At3g12630, At3g23590, At3g61140, At3g62130, At3g62010, At3g63520, At4g15530*, and *At5g49480* were identified as potential candidates. *At3g06050* encodes ATPRXIIF, which is involved in oxidative stress response. ATPRXIIF is a mitochondrial thioredoxin required for the proper functioning of major metabolic pathways, including stomatal function and antioxidant metabolism under sub-optimal conditions such as drought and salinity (Geigenberger et al., 2017). Furthermore, the inactivation of the mitochondrial TRX system leads to metabolite adjustments in both primary and secondary metabolism following drought episodes in Arabidopsis, making the plants more resistant to drought stress (da Fonseca-Pereira et al., 2019). As this is a mitochondrial trait, IL 1-2 should be considered a maternal donor for use in breeding programs. *At3g12630* encodes SAP5. In wheat (*Triticum aestivum*) and *Arabidopsis thaliana*, STRESS-SAP5 is involved in drought tolerance and acts as an E3 ubiquitin ligase to target DRIP and MBP-1 for degradation (Zhang et al., 2019). *At3g23590* encodes REF4-RELATED 1, which is involved in the regulation of the phenylpropanoid metabolic process (Bonawitz et al., 2012). The phenylpropanoid biosynthetic pathway is activated under abiotic stress conditions such as drought, resulting in the accumulation of various phenolic compounds, which, among other roles, scavenge harmful ROS (Sharma et al., 2019). Li et al. (2020) found that an elevated phenylpropanoid content enhanced wheat tolerance to drought stress by maintaining an improved water status, leading to higher photosynthesis rates and lower membrane damage. The gene *At3g62130* encodes LCD. Hydrogen sulfide (H_2S) plays a crucial role in the regulation of stomatal closure in plant responses to drought stress, and LCD has been identified as being primarily responsible for the degradation of cysteine to generate H_2S (Jin et al., 2011). H_2S may be an essential link in the stomatal regulation by ABA via ion channels. H_2S affects the expression of ABA receptor candidates, and ABA also influences H_2S production. Thus, H_2S interacts with ABA in stomatal regulation, which is responsible for drought stress tolerance in *Arabidopsis* (Jin et al., 2013). In addition, according to Zhou et al. (2021), the overexpression of LCD results in enhanced H_2S production and the persulfidation of total soluble protein, and confers drought tolerance in rice. *At3g63520* encodes 9-cis-epoxycarotenoid dioxygenase, which is involved in the response to water deprivation. This enzyme is thought to be a key enzyme in ABA biosynthesis (Iuchi et al., 2000). ABA is a plant hormone involved in stress responses, and is quickly accumulated by many plant species when exposed to drought stress (Lim et al., 2015). Iuchi et al. (2001) linked 9-cis-epoxycarotenoid dioxygenase and the response to water deprivation by demonstrating that

drought induces the expression of the gene encoding this enzyme, which in turn controls the level of endogenous ABA under drought-stressed conditions.

Uncovering genes in IL populations is a powerful method of mining the genetic variation present in wild crop relatives for traits with the potential to improve desired characteristics in elite varieties. Several traits identified in ILs and advanced background populations have been used in tomato breeding (Lippman et al., 2007). In this study, we identified several candidate genes related to drought tolerance in ILs. We provide a valuable aid for breeding programs with useful information for classical breeding approaches by indicating promising genotypes to be used as genitors, and for molecular breeding approaches by identifying candidate genes that can be introgressed into elite varieties through marker-assisted selection. Furthermore, our results represent a significant step toward developing drought-tolerant tomato cultivars.

5. CONCLUSION

The FAI-BLUP index ranked IL 1-4-18, IL 7-4-1, IL 7-1, IL 7-5-5, and IL 1-2 as closest to the drought-tolerant ideotype. IL 1-4-18 and IL1-2 were also previously identified as drought-tolerant during the germination and seedling stages.The following candidate genes associated with drought tolerance were identified in selected ILs: *At1g55870, At1g55840, At3g06050, At3g12630, At3g23590, At3g26060, At3g43230, At3g62130, At3g63520,* and *At4g38810*. Because they are already associated with genetic markers, they can be transferred to elite tomato cultivars through marker-assisted technology after validation

REFERENCES

Ahammed, G. J., Li, X., Yang, Y., Liu, C., Zhou, G., Wan, H., & Cheng, Y. (2020). Tomato WRKY81 acts as a negative regulator for drought tolerance by modulating guard cell H2O2–mediated stomatal closure. *Environmental and Experimental Botany*, *171*, 103960. https://doi.org/10.1016/J.ENVEXPBOT.2019.103960

Ayankojo, I. T., Morgan, K. T., Ozores-Hampton, M., & Migliaccio, K. W. (2018). Effects of real-time location-specific drip irrigation scheduling on water use, plant growth, nutrient accumulation, and yield of Florida fresh-market tomato. *HortScience*, *53*(9), 1372–1378. https://doi.org/10.21273/HORTSCI13183-18

Bartlett, M. S., 1937. The statistical conception of mental factors. British Journal of Psychology, 28, 97–104. https://doi.org/10.1111/j.2044-8295.1937.tb00863.x.

Bernardo, S., Soares, A. A., & Mantovani, E. C. 2019.Manual de irrigação. Editora UFV.

Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., Quesneville, H., Alseekh, S., Sørensen, I., Lichtenstein, G., Fich, E. A., Conte, M., Keller, H., Schneeberger, K., Schwacke, R., Ofner, I., Vrebalov, J., Xu, Y., Osorio, S., … Fernie, A. R. (2014). The genome of the stress-tolerant wild tomato species Solanum pennellii. *Nature Genetics*, *46*(9), 1034–1038. https://doi.org/10.1038/ng.3046

Bonawitz, N. D., Soltau, W. L., Blatchley, M. R., Powers, B. L., Hurlock, A. K., Seals, L. A., Weng, J. K., Stout, J., & Chapple, C. (2012). REF4 and RFR1, subunits of the transcriptional coregulatory complex mediator, are required for phenylpropanoid homeostasis in Arabidopsis. *Journal of Biological Chemistry*, *287*(8), 5434–5445. https://doi.org/10.1074/JBC.M111.312298/ATTACHMENT/700807FE-CE97-40EF-B4DB-1834998F07FE/MMC1.PDF

Covarrubias, A. A., & Reyes, J. L. (2010). Post-transcriptional gene regulation of salinity and drought responses by plant microRNAs. *Plant, Cell & Environment*, *33*(4), 481–489. https://doi.org/10.1111/J.1365-3040.2009.02048.X

Cui, J., Shao, G., Lu, J., Keabetswe, L., & Hoogenboom, G. (2019). Yield, quality and drought sensitivity of tomato to water deficit during different growth stages. *Scientia Agricola*, *77*(2), 2020. https://doi.org/10.1590/1678-992X-2018-0390

da Fonseca-Pereira, P., Daloso, D. M., Gago, J., Nunes-Nesi, A., & Araújo, W. L. (2019). On the role of the plant mitochondrial thioredoxin system during abiotic stress. *Https://Doi.Org/10.1080/15592324.2019.1592536*, *14*(6). https://doi.org/10.1080/15592324.2019.1592536

Dariva, F. D., Copati, M. G. F., Pessoa, H. P., Alves, F. M., Dias, F. de O., Picoli, E. A. de T., da Cunha, F. F., & 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 2020 10:1*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-73004-4

Daszkowska-Golec, A. (2016). The Role of Abscisic Acid in Drought Stress: How ABA Helps Plants to Cope with Drought Stress. *Drought Stress Tolerance in Plants, Vol 2: Molecular and Genetic Perspectives*, 123–151. https://doi.org/10.1007/978-3-319-32423-4_5

Ebert, A. W., & Schafleitner, R. (2015). *Utilization of Wild Relatives in the Breeding of Tomato and Other Major Vegetables*.

Egea, I., Albaladejo, I., Meco, V., Morales, B., Sevilla, A., Bolarin, M. C., & Flores, F. B. (2018). *The drought-tolerant Solanum pennellii regulates leaf water loss and induces genes involved in amino acid and ethylene/jasmonate metabolism under dehydration OPEN*. *8*, 2791. https://doi.org/10.1038/s41598-018-21187-2

Eshed, Y., & Zamir, D. (1995). An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield- associated QTL. *Genetics*, *141*(3), 1147–1162.

FAO (2021). Food and Agriculture Organization of the Union. https://doi.org/10.4060/cb2395fr (Consulted on 10/02/2022).

Florido Bacallao, M., & Bao Fundora, L. (2014). Revisión bibliográfica.Tolerancia a estrés por déficit hídrico en tomate (Solanum lycopersicum L.). *Cultivos Tropicales*, *35*(3), 70–88. http://ediciones.inca.edu.cu

Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, *2007*. https://doi.org/10.1155/2007/64358

Foolad, M. R., & Foolad, M. R. (2004). Recent advances in genetics of salt tolerance in Tomato. *Plant Cell, Tissue and Organ Culture*, *76*(2), 101–119. https://doi.org/10.1023/B:TICU.0000007308.47608.88

Foolad, M. R., Subbiah, P., & Zhang, L. (2007). Common QTL affect the rate of tomato seed germination under different stress and nonstress conditions. *International Journal of Plant Genomics*, *2007*. https://doi.org/10.1155/2007/97386

Fujita, Y., Nakashima, K., Yoshida, T., Katagiri, T., Kidokoro, S., Kanamori, N., Umezawa, T., Fujita, M., Maruyama, K., Ishiyama, K., Kobayashi, M., Nakasone, S., Yamada, K., Ito, T., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2009). Three SnRK2 Protein Kinases are the Main Positive Regulators of Abscisic Acid Signaling in Response to Water Stress in Arabidopsis. *Plant and Cell Physiology*, *50*(12), 2123–2132. https://doi.org/10.1093/PCP/PCP147

Geigenberger, P., Thormählen, I., Daloso, D. M., & Fernie, A. R. (2017). The Unprecedented Versatility of the Plant Thioredoxin System. *Trends in Plant Science*, *22*(3), 249–262. https://doi.org/10.1016/J.TPLANTS.2016.12.008

Genuchten, M. Th. van. (1980). A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/SSSAJ1980.03615995004400050002X

George, S., Ahmad Jatoi, S., & Uddin Siddiqui, S. (2013). Genotypic differences against PEG simulated drought stress in tomato. *Pakistan Journal of Botany*, *45*(5), 1551–1556.

Goyal, M. R., & Sharma, P. (2018). Water Requirement of Drip-Irrigated Tomato Inside a Shade Net House. In *Engineering Interventions in Sustainable Trickle Irrigation* (pp. 35–42). Apple Academic Press. https://doi.org/10.1201/9781315184241-2

Gur, A., Semel, Y., Osorio, S., Friedmann, M., Seekh, S., Ghareeb, B., Mohammad, A., Pleban, T., Gera, G., Fernie, A. R., & Zamir, D. (2011). Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. *Theoretical and Applied Genetics*, *122*(2), 405–420. https://doi.org/10.1007/S00122-010-1456-9/FIGURES/6

Henderson, C. R. (1975). *Best Linear Unbiased Estimation and Prediction under a Selection Model* (Vol. 31, Issue 2).

Huang, J., Ghosh, R., & Bankaitis, V. A. (2016). Sec14-like phosphatidylinositol transfer proteins and the biological landscape of phosphoinositide signaling in plants. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1861*(9), 1352–1364. https://doi.org/10.1016/J.BBALIP.2016.03.027

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., Tabata, S., Kakubari, Y., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2001). Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. *The Plant Journal*, *27*(4), 325–333. https://doi.org/10.1046/J.1365-313X.2001.01096.X

Iuchi, S., Kobayashi, M., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2000). A Stress-Inducible Gene for 9-cis-Epoxycarotenoid Dioxygenase Involved in Abscisic Acid

Biosynthesis under Water Stress in Drought-Tolerant Cowpea. *Plant Physiology*, *123*(2), 553–562. https://doi.org/10.1104/PP.123.2.553

Jin, Z., Shen, J., Qiao, Z., Yang, G., Wang, R., & Pei, Y. (2011). Hydrogen sulfide improves drought resistance in Arabidopsis thaliana. *Biochemical and Biophysical Research Communications*, *414*(3), 481–486. https://doi.org/10.1016/J.BBRC.2011.09.090

Jin, Z., Xue, S., Luo, Y., Tian, B., Fang, H., Li, H., & Pei, Y. (2013). Hydrogen sulfide interacting with abscisic acid in stomatal regulation responses to drought stress in Arabidopsis. *Plant Physiology and Biochemistry*, *62*, 41–46. https://doi.org/10.1016/J.PLAPHY.2012.10.017

Jing, L. W., Chen, S. H., Guo, X. L., Zhang, H., & Zhao, Y. X. (2006). Overexpression of a Chloroplast-located Peroxiredoxin Q Gene, SsPrxQ, Increases the Salt and Low-temperature Tolerance of Arabidopsis. *Journal of Integrative Plant Biology*, *48*(10), 1244–1249. https://doi.org/10.1111/J.1744-7909.2006.00357.X

Kang, M., Fokar, M., Abdelmageed, H., & Allen, R. D. (2011). Arabidopsis SAP5 functions as a positive regulator of stress responses and exhibits E3 ubiquitin ligase activity. *Plant Molecular Biology*, *75*(4–5), 451–466. https://doi.org/10.1007/S11103-011-9748-2/FIGURES/8

Kulik, A., Wawer, I., Krzywińska, E., Bucholc, M., & Dobrowolska, G. (2011). SnRK2 protein Kinases - Key regulators of plant response to abiotic stresses. *OMICS A Journal of Integrative Biology*, *15*(12), 859–872. https://doi.org/10.1089/OMI.2011.0091/ASSET/IMAGES/LARGE/FIGURE3.JPEG

kumar, P. P., Acharya, B., Dutta, S., Saha, S., & Hazra, P. (2019). Variation among the genotypes of tomato for ceding and in-vitro culture for increasrossability to its wild relatives . *International Journal of Chemical Studies*, *7*(3), 3423–3426. https://www.researchgate.net/publication/334987700

Li, X., Zhang, X., Liu, G., Tang, Y., Zhou, C., Zhang, L., & Lv, J. (2020). The spike plays important roles in the drought tolerance as compared to the flag leaf through the phenylpropanoid pathway in wheat. *Plant Physiology and Biochemistry*, *152*, 100–111. https://doi.org/10.1016/J.PLAPHY.2020.05.002

Liebthal, M., Maynard, D., & Dietz, K. J. (2018). Proxiredoxins and Redox Signaling in Plants. *Antioxidants and Redox Signaling*, *28*(7), 609–624. https://doi.org/10.1089/ARS.2017.7164/ASSET/IMAGES/LARGE/FIGURE7.JPEG

Lim, C. W., Baek, W., Jung, J., Kim, J. H., & Lee, S. C. (2015). Function of ABA in Stomatal Defense against Biotic and Drought Stresses. *International Journal of Molecular Sciences 2015, Vol. 16, Pages 15251-15270*, *16*(7), 15251–15270. https://doi.org/10.3390/IJMS160715251

Lippman, Z. B., Semel, Y., & Zamir, D. (2007a). An integrated view of quantitative trait variation using tomato interspecific introgression lines. In *Current Opinion in Genetics and Development* (Vol. 17, Issue 6, pp. 545–552). https://doi.org/10.1016/j.gde.2007.07.007

Lippman, Z. B., Semel, Y., & Zamir, D. (2007b). An integrated view of quantitative trait variation using tomato interspecific introgression lines. In *Current Opinion in Genetics and Development* (Vol. 17, Issue 6, pp. 545–552). Elsevier Current Trends. https://doi.org/10.1016/j.gde.2007.07.007

López-Galiano, M. J., García-Robles, I., González-Hernández, A. I., Camañes, G., Vicedo, B., Real, M. D., & Rausell, C. (2019). Expression of miR159 Is Altered in Tomato Plants Undergoing Drought Stress. *Plants 2019, Vol. 8, Page 201*, *8*(7), 201. https://doi.org/10.3390/PLANTS8070201

Makalesi, A. (2020). Effectiveness of In Vitro and In Vivo Tests for Screening of Tomato Genotypes against Drought Stress. *Ege Üniv. Ziraat Fak. Derg*, 143–150. https://doi.org/10.20289/zfdergi.835798

Mangat, P. K., Shim, J., Gannaban, R. B., Singleton, J. J., & Angeles-Shim, R. B. (2021). Alien introgression and morpho-agronomic characterization of diploid progenies of Solanum lycopersicoides monosomic alien addition lines (MAALs) toward pre-breeding applications in tomato (S. lycopersicum). *Theoretical and Applied Genetics*, *134*(4), 1133–1146. https://doi.org/10.1007/S00122-020-03758-Y/FIGURES/6

Mitchell, J. P., Shennan, C., & Grattan, S. R. (1991). Developmental changes in tomato fruit composition in response to water deficit and salinity. *Physiologia Plantarum*, *83*(1), 177–185. https://doi.org/10.1111/J.1399-3054.1991.TB01299.X

Monks, D. E., Aghoram, K., Courtney, P. D., DeWald, D. B., & Dewey, R. E. (2001). Hyperosmotic Stress Induces the Rapid Phosphorylation of a Soybean Phosphatidylinositol Transfer Protein Homolog through Activation of the Protein Kinases SPK1 and SPK2. *The Plant Cell*, *13*(5), 1205–1219. https://doi.org/10.1105/TPC.13.5.1205

Moyle, L. C., & 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

Pashkovskiy, P. P., Vankova, R., Zlobin, I. E., Dobrev, P., Ivanov, Y. v., Kartashov, A. v., & Kuznetsov, V. v. (2019). Comparative analysis of abscisic acid levels and expression of abscisic acid-related genes in Scots pine and Norway spruce seedlings under water deficit. *Plant Physiology and Biochemistry*, *140*, 105–112. https://doi.org/10.1016/J.PLAPHY.2019.04.037

Patterson, H. D., & Thompson, R. (1971). Recovery of Inter-Block Information when Block Sizes are Unequal. *Biometrika*, *58*(3), 545. https://doi.org/10.2307/2334389

Pessoa, H. P. Linhagens de introgressão derivadas de *Solanum pennellii*: tolerância ao déficit hídrico durante a germinação e no estádio de plântula. MSc dissertation, Universidade Federal de Viçosa, Departamento de Fitotecnia, (2018).

Richter, H. (1997). Water relations of plants in the field: some comments on the measurement of selected parameters. *Journal of Experimental Botany*, *48*(1), 1–7. https://doi.org/10.1093/JXB/48.1.1

Ritchie, G. A., & Hinckley, T. M. (1975). The Pressure Chamber as an Instrument for Ecological Research. *Advances in Ecological Research*, *9*(C), 165–254. https://doi.org/10.1016/S0065-2504(08)60290-1

Rocha, J. R. do A. S. de C., Machado, J. C., & Carneiro, P. C. S. (2018). Multitrait index based on factor analysis and ideotype-design: proposal and application on elephant grass breeding for bioenergy. *GCB Bioenergy*, *10*(1), 52–60. https://doi.org/10.1111/gcbb.12443

Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Bruedigam, C., Kopka, J., Willmitzer, L., Zamir, D., & Fernie, A. R. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology 2006 24:4*, *24*(4), 447–454. https://doi.org/10.1038/nbt1192

Scholander, P. F., Hammel, H. T., Bradstreet, E. D., & Hemmingsen, E. A. (1965). Sap Pressure in Vascular Plants. *Science*, *148*(3668), 339–346. https://doi.org/10.1126/SCIENCE.148.3668.339

Seki, K. (2007). SWRC fit – a nonlinear fitting program with a water retention curve for soils having unimodal and bimodal pore structure. *Hydrol. Earth Syst. Sci. Discuss*, *4*, 407–437.

Sellin, A. (1996). Base water potential of Picea abies as a characteristic of the soil water status. *Plant and Soil 1996 184:2*, *184*(2), 273–280. https://doi.org/10.1007/BF00010456

Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., & Zheng, B. (2019a). Response of Phenylpropanoid Pathway and the Role of Polyphenols in Plants under Abiotic Stress. *Molecules (Basel, Switzerland)*, *24*(13). https://doi.org/10.3390/MOLECULES24132452

Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., & Zheng, B. (2019b). Response of Phenylpropanoid Pathway and the Role of Polyphenols in Plants under Abiotic Stress. *Molecules 2019, Vol. 24, Page 2452*, *24*(13), 2452. https://doi.org/10.3390/MOLECULES24132452

Suprunova, T., Krugman, T., Distelfeld, A., Fahima, T., Nevo, E., & Korol, A. (2007). Identification of a novel gene (Hsdr4) involved in water-stress tolerance in wild barley. *Plant Molecular Biology*, *64*(1–2), 17–34. https://doi.org/10.1007/S11103-006-9131-X/FIGURES/8

Toubiana, D., Semel, Y., Tohge, T., Beleggia, R., Cattivelli, L., Rosental, L., Nikoloski, Z., Zamir, D., Fernie, A. R., & Fait, A. (2012). Metabolic profiling of a mapping population exposes new insights in the regulation of seed metabolism and seed, fruit, and plant relations. *PLoS Genetics*, *8*(3). https://doi.org/10.1371/journal.pgen.1002612

Vanisree, P., Djanaguiraman, M., Sankari, A., & Pugalendhi, L. (2021). Effect of deficit irrigation on physiological and biochemical traits of tomato (Solanum lycopersicum) genotypes. *The Pharma Innovation Journal*, *10*(10), 2134–2140. http://www.thepharmajournal.com

Wudiri, B. B., & Henderson, D. W. (1985). Effects of water stress on flowering and fruit set in processing-tomatoes. *Scientia Horticulturae*, *27*, 189–198.

Yang, H., Du, T., Qiu, R., Chen, J., Wang, F., Li, Y., Wang, C., Gao, L., & Kang, S. (2017). Improved water use efficiency and fruit quality of greenhouse crops under regulated deficit irrigation in northwest China. *Agricultural Water Management*, *179*, 193–204. https://doi.org/10.1016/J.AGWAT.2016.05.029

Zang, D., Wang, C., Ji, X., & Wang, Y. (2015). Tamarix hispida zinc finger protein ThZFP1 participates in salt and osmotic stress tolerance by increasing proline content and SOD and POD activities. *Plant Science*, *235*, 111–121. https://doi.org/10.1016/J.PLANTSCI.2015.02.016

Zhang, J. F., Yuan, L. J., Shao, Y., Du, W., Yan, D. W., & Lu, Y. T. (2008). The disturbance of small RNA pathways enhanced abscisic acid response and multiple stress responses in Arabidopsis. *Plant, Cell & Environment*, *31*(4), 562–574. https://doi.org/10.1111/J.1365-3040.2008.01786.X

Zhang, N., Xu, J., Liu, X., Liang, W., Xin, M., Du, J., Hu, Z., Peng, H., Guo, W., Ni, Z., Sun, Q., & Yao, Y. (2019). Identification of HSP90C as a substrate of E3 ligase TaSAP5 through ubiquitylome profiling. *Plant Science*, *287*, 110170. https://doi.org/10.1016/J.PLANTSCI.2019.110170

Zhao, T., Wu, T., Pei, T., Wang, Z., Yang, H., Jiang, J., Zhang, H., Chen, X., Li, J., & Xu, X. (2021). Overexpression of SlGATA17 Promotes Drought Tolerance in Transgenic Tomato Plants by Enhancing Activation of the Phenylpropanoid Biosynthetic Pathway. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/FPLS.2021.634888/FULL

Zhou, H., Zhou, Y., Zhang, F., Guan, W., Su, Y., Yuan, X., & Xie, Y. (2021). Persulfidation of Nitrate Reductase 2 Is Involved in l-Cysteine Desulfhydrase-Regulated Rice Drought Tolerance. *International Journal of Molecular Sciences 2021, Vol. 22, Page 12119*, *22*(22), 12119. https://doi.org/10.3390/IJMS2222121194

Zsögön, A., Vicente, M. H., Reartes, D. S., & Peres, L. E. P. (2017). Understanding and improving water-use efficiency and drought resistance in tomato. In A. Mattoo & A. Handa (Eds.), *Achieving sustainable cultivation of tomatoes* (1st ed., pp. 117–142). Burleigh Dodds Science Publishing. https://doi.org/10.19103/AS.2016.0007.01

OVERALL CONCLUSIONS

Based on our results, we can affirm that the *S. pennellii* IL population is a valuable source of gene reservoirs for seed quality and drought tolerance. IL 6-1, 4-1-1, 4-4, IL 11-4-1, IL 1-1-2, and IL 10-1 were the most similar to the ideotype of seed germination performance and seedling vigor. IL 1-4-18 and IL1-2 were drought-tolerant during both the seedling and vegetative/reproductive stages. Therefore, these genotypes should be included in breeding programs.

The genes *At1g75350, At2g26350, At2g28490, At3g10920, At5g13200, At5g41480,* and At4g21800 were candidate genes for seed germination. The genes *At5g53000, At1g46480, At1g30755, At5g47390*, *At2g29630*, and *At4g03400* were candidate genes for seedling vigor. The genes *At3g03790, At3g06050, At3g12630, At3g15530, At3g23590, At3g49480, At3g61140, At3g62130, At3g62010,* and *At3g63520* were candidate genes for drought tolerance. As these genes are already associated with genetic markers, they can be transferred to elite tomato cultivars through marker-assisted technology, directly after validation.