

**University of São Paulo
“Luiz de Queiroz” College of Agriculture**

**Genomic dissection of a tropical maize diversity panel: a study on molecular
characterization and resistance to the corn stunt disease complex**

Fernando Garcia Espolador

Thesis presented to obtain the degree of Doctor in Science.
Area: Genetics and Plant Breeding

**Piracicaba
2020**

Fernando Garcia Espolador
Agronomic Engineer

**Genomic dissection of a tropical maize diversity panel: a study on molecular
characterization and resistance to the corn stunt disease complex**

Advisor:
Prof. Dr. **ROBERTO FRITSCHÉ NETO**

Thesis presented to obtain the degree of Doctor in Science.
Area: Genetics and Plant Breeding

Piracicaba
2020

**To my parents, Carlos and Maria de Fátima,
and my fiancée, Júlia,
who have always supported me**

ACKNOWLEDGMENTS

I would like to address my acknowledgments:

To God, for all the blessings, the wisdom, and for guiding my steps on my journey.

To my parents, Carlos Maschio Espolador and Maria de Fátima Garcia Espolador, for their love and support, for being examples for me and the source of my personal values.

To my brother, Murilo Garcia Espolador, for his friendship.

To my fiancée, Julia Silva Morosini, for her love, support, confidence, strength, and for being an example of professional and person to me.

To my parents-in-law, Carlos Alberto Morosini and Olivia Maria B. S. Morosini, and my sister-in-law, Natalia Silva Morosini, for their constant support and friendship.

To the University of São Paulo and College of Agriculture 'Luiz de Queiroz', for the opportunities and for my professional qualification.

To Prof. Dr. Roberto Fritsche Neto and the Allogamous team, for the opportunity to work with maize breeding and for the lessons.

To the University of Wisconsin-Madison, for receiving me as exchange student.

To Prof. Dr. Natalia de Leon, for the opportunity of internship abroad, her confidence, knowledge and for being an example of professionalism.

To the funding agencies. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, National Council for Scientific and Technological Development (CNPq), and São Paulo Research Foundation – FAPESP (Grant 2017/24327-0).

Finally, to all those who directly or indirectly helped in this work, in my personal training and professional qualification, my sincere THANK YOU!

RESUMO

Análise genômica de um painel de diversidade de milho tropical: um estudo sobre caracterização molecular e resistência ao complexo do enfezamento do milho

O milho tropical é uma importante fonte de diversidade genética a ser explorada por programas de melhoramento para lidar com novos desafios agrícolas, mas isto demanda caracterização genética. Por exemplo, o complexo do enfezamento do milho (enfezamento pálido e enfezamento vermelho) tem causado perdas significativas de produtividade nas últimas safras, contudo o controle genético e o germoplasma resistente para estas doenças são pouco compreendidos. Neste trabalho, um painel de diversidade de milho tropical foi construído e caracterizado geneticamente visando delinear um pool representativo de germoplasma tropical para estudos genéticos e investigar os genes associados à resistência do complexo do enfezamento do milho, assim como as potenciais fontes de resistência. Para isso, 360 linhagens altamente diversas geneticamente foram genotipadas usando a abordagem de genotipagem por sequenciamento com as enzimas de restrição PstI e MseI e alinhadas com a versão 5 do genoma de referência B73. Para as avaliações genéticas, foram considerados dois conjuntos de dados: um com dados brutos e o segundo com dados imputados e filtrados para controle de qualidade, mantendo apenas marcadores bialélicos com frequência alelo raro maior que 0,05, valor de chamada de SNPs maior que 0,95 e desequilíbrio de ligação (r^2) menor que 0,99. O estudo de associação ampla do genoma (GWAS) foi realizado incorporando a matriz de relacionamento genômico e três componentes principais para lidar com a estrutura do painel. Usando o conjunto de dados filtrado, foram analisadas as características proporção de plantas sobreviventes (PSP), nota de sanidade das plantas sobreviventes (SSSP) e nota de sanidade total (WSS), cujas mensurações ocorreram em dois locais no estado de São Paulo, Brasil. O conjunto de dados brutos incluiu 196.803 SNPs bem distribuídos pelos cromossomos. A proporção de dados faltantes foi de 0,391 e a heterozigosidade média observada foi de 0,036. O conjunto de dados filtrados, contendo 14.655 SNPs, apresentou estimativas semelhantes de parâmetros genéticos populacionais em comparação ao primeiro conjunto. A análise da estrutura do painel apontou a existência de nove subpopulações. Através da GWAS, 13 marcadores foram significativos para as características avaliadas e apresentaram funções relacionadas principalmente ao metabolismo da celulose, via da auxina, genes de defesa e resposta associados à atividade de fagócito oxidase e produção de antocianinas. Os genes candidatos foram associados, por exemplo, à resposta do acúmulo de glicose nas folhas, redução do teor de auxina na planta e ataque direto aos patógenos. Para cada característica, foi possível identificar linhagens com todos os alelos favoráveis em homozigose, o que facilitaria a transferência de genes de resistência para outros genótipos. Além disso, as linhagens advindas do ancestral PF-41X05-33-05B apresentaram alta concentração de alelos favoráveis a todas as características simultaneamente. Os presentes resultados revelam mecanismos genéticos subjacentes desencadeados por plantas em resposta ao complexo do enfezamento do milho e permitiram a identificação de potenciais linhagens resistentes. Tais inferências podem melhorar substancialmente os ganhos genéticos quando incorporadas em programas de melhoramento e constituem uma importante contribuição para a compreensão genética do germoplasma de milho tropical.

Palavras-chave: *Zea mays*, Painel de diversidade, *Spiroplasma kunkelii*, Maize bushy stunt phytoplasma

ABSTRACT

Genomic dissection of a tropical maize diversity panel: a study on molecular characterization and resistance to the corn stunt disease complex

The tropical maize is an important pool of genetic diversity to be explored in breeding programs to face new agricultural challenges, nevertheless it requires genetic characterization. For instance, the corn stunt disease complex (corn stunt disease and maize bushy stunt disease) caused significant losses in the last crop seasons, but its genetic control and resistant germplasm are poorly comprehended. In this study, we assembled and genetically characterized a tropical maize diversity panel in order to construct a representative pool of tropical germplasm for genetic studies and to investigate the genes associated with the resistance of corn stunt disease complex, as well as the potential sources of resistance. For that, 360 inbred lines highly diverse were genotyped using a genotyping-by-sequencing approach with restriction enzymes PstI and MseI and aligned to the version 5 of the B73 reference genome. For genetic assessments, two datasets were considered: one with the raw data, and the second one with data imputed and filtered for quality control, retaining only biallelic markers with minor allelic frequency higher than 0.05, call rate higher than 0.95, and linkage disequilibrium (r^2) lower than 0.99. Genome wide association study (GWAS) was performed incorporating the genomic relationship matrix and 3 main principal components to deal with panel structure. Using the filtered dataset, we analyzed the traits proportion of survivor plants (PSP), sanity score in survivor plants (SSSP), and whole sanity score (WSS) which were evaluated in two sites in the State of São Paulo, Brazil. The raw dataset contained 196,803 SNPs well distributed across the chromosomes. The proportion of missing data was 0.391 and the average observed heterozygosity was 0.036. The filtered dataset, containing 14,655 SNPs, showed similar estimates of populational genetic parameters compared to the first one. The structure analysis indicated that this panel comprises nine subpopulations. Through GWAS, 13 markers were significant to the traits and presented functions mainly related to cellulose metabolism, auxin pathway, genes of defense and response to phagocyte oxidase activity, and anthocyanin production. The candidate genes were associated, for example, with the response of glucose accumulation in leaves, the reduction of auxin content, and the direct defense-attack against the pathogens. For each of the traits, we found lines with the totality of favorable alleles in homozygosity, which would facilitate transferring resistance genes to other genotypes. Additionally, the lines from the ancestral PF-41X05-33-05B exhibited high content of favorable alleles to all the traits simultaneously. Our findings reveal underlying genetic mechanisms triggered by plants in response to the corn stunt disease complex and allowed the identification of potential resistant inbred lines. These results can substantially improve the genetic gains when incorporated in breeding programs and constitute an important contribution to the genetic comprehension of the tropical maize germplasm.

Keywords: *Zea mays*, Diversity panel, Structure, *Spiroplasma kunkelii*, Maize bushy stunt phytoplasma

1. INTRODUCTION

Maize (*Zea mays*) is a crop domesticated from its ancestral species teosinte (*Zea teosinte*) by ancient civilizations near Rio Balsas region, Mexico, 8700 years ago (Mir et al., 2013). The natives domesticated this crop by selecting plants that presented only one stalk, bigger ears with more grains, absence of shell around kernels, and indehiscence (Doebley, 2004; Szabó & Burr, 1996). As result of such intense selection process, the derived species, *Zea mays*, had a wide spread across the pre-Columbian societies in America, leading to its dispersion to different environments: from low to high altitudes, ranging from temperate regions of North America, passing by Central American plains until areas in high altitude in Andes (Edmeades et al., 2017; Matsuoka et al., 2002). Consequently, the maize germplasm presents adaptation to cultivation in regions at up to 4000 m of altitude, in latitudes from 0 to 57°, on acid to basic soils, and in a broad range of annual rainfalls regimes (400–2500 mm) (Edmeades et al., 2017).

When the European arrived in the American continent and discovered maize, they dispersed it around the Europe and the world (Mir et al., 2013). Nowadays, this is the cereal most cultivated worldwide, with annual production of 1.45 billion tons of grain (FAOSTAT, 2020; Leff et al., 2004). This cereal is used as an essential food staple for several societies, raw material for industry, feed for livestock, and as renewable biofuel (Ranum et al., 2014).

Despite the great variability presented in the species, the germplasm pools are not fully explored. For instance, Hallauer & Miranda (1988) reported the existence of 130 distinct races in maize germplasm, from which 71% were originated in South America according Serratos-Hernández (2009). However, this diversity is underexplored, since studies on tropical maize are scarce and comprehend only part of the genetic variability existent (Laborda et al., 2005; Lanés et al., 2014; Wu et al., 2016). The exploitation of the tropical germplasm can contribute to meet current and future demands, to identify hybrids adapted to marginal crop conditions, and to find sources of resistance to abiotic and biotic stresses.

The temperate maize presents a better characterization, heterotic groups clearer defined and higher potential yield compared to the tropical maize (Edmeades et al., 2017; Stevenson & Goodman, 1972). Conversely, the latter presents a greater genetic variability.

Phytopathogenic mollicutes represent a threat to the current maize production, especially *Spiroplasma kunkelii* and the *maize bush stunt phytoplasma* (MBSP, also known as *Candidatus Phytoplasma asteris*). They are responsible for the corn stunt disease and maize bush stunt disease, respectively, and together are called the corn stunt disease complex (Bergamin Filho et al., 1995; Kimati et al., 2005). The disease complex is spread across the American continent and is transmitted by a hemipteran vector, e.g. *Dalbulus maidis*. Previously considered a secondary disease, it became greatly

important with the changes in the production system, such as the cultivation in a fall-winter second season ('safrinha'), leading to impressive losses (Kimati et al., 2005).

The exploitation of the genetic variability that exists in tropical maize for this set of diseases can lead to a better comprehension of the pathosystem and be a valuable resource for farmers and plant breeding programs. In this context, genome wide association studies (GWAS) can be applied to identify resistant genes associated with the complex, to assist understanding the metabolic pathways involved, and to find genotypes containing high proportion of favorable alleles.

Given that the construction of a representative diversity panel of tropical maize is a powerful tool to comprehend the genetic variability available, to discover genes and genotypes to improve the tropical agriculture and to attend future human and environmental demands, this study aimed to assemble and genotypically characterize a tropical maize diversity panel and to investigate the genes associated with the resistance of corn stunt disease complex, as well as the potential sources of resistance.

2. CONCLUSION

The tropical maize diversity panel evaluated in this study presented a considerable genetic variability. The filtered data yielded estimations similarly to the raw dataset, but with reduction in the population effective size due to the exclusion of the non-biallelic markers and the diversity present in rare alleles. The structure analysis designated nine groups as the most parsimonious clustering of the panel. They provided a consistent classification of genotypes according to their origin. The similarity of genetic parameters obtained from this panel and from current commercial germplasm indicates a substantial resemblance between them, which must be explored better in future studies. It was possible to identify genetic variability for the traits PSP, SSSP and WSS concerning the corn stunt disease and bushy maize stunt disease in this tropical maize diversity panel. In total, we found 13 markers associated to the disease complex, and the candidate genes identified support and elucidates the metabolic pathways affected. The genes were associated mainly to auxin metabolism, cellulose biosynthesis and defense-protection against pathogens genes. The markers identified were well distributed across the germplasm, being possible to identify lines with high content of favorable alleles for each trait. Inbred lines descendent of PF-41X05-33-05B presented high concentration of favorable alleles for the three traits simultaneously. Therefore, they may be used directly to produce hybrids or incorporated in recurrent-selection programs to enhance the favorable alleles in the population. Future researches must validate and verify the occurrence of non-additive effects. Therefore, these finds play an important role in enhancing the knowledge about tropical maize germplasm and assisting breeding programs.

REFERENCES

- Bergamin Filho, A., Kimati, H., & Amorim, L. (1995). Manual de Fitopatologia: Princípios e Conceitos (A. Bergamin Filho, H. Kimati, & L. Amorim (eds.); 3rd ed.). Agronômica Ceres.
- Doebley, J. (2004). The Genetics of Maize Evolution. *Annual Review of Genetics*, 38(1), 37–59. <https://doi.org/10.1146/annurev.genet.38.072902.092425>
- Edmeades, G. O., Trevisan, W., Prasanna, B. M., & Campos, H. (2017). Tropical Maize (*Zea mays* L.). In H. Campos & P. D. S. Caligari (Eds.), *Genetic improvement of tropical crops* (pp. 57–109). <https://doi.org/10.1007/978-3-319-59819-2>
- FAOSTAT. (2020). FAOSTAT. <http://www.fao.org/faostat/en/?#data/QC>
- Hallauer, A. R., & Miranda, J. B. (1988). *Quantitative Genetics in Maize Breeding* (2nd ed.). Iowa State University Press.
- Kimati, H., Amorim, L., Rezende, J. A. M., Bergamin Filho, A., & Camargo, L. E. A. (2005). Manual de Fitopatologia, Vol II. In *Doenças das plantas cultivadas*.
- Laborda, P. R., Oliveira, K. M., Garcia, A. A. F., Paterniani, M. E. A. G. Z., & De Souza, A. P. (2005). Tropical maize germplasm: What can we say about its genetic diversity in the light of molecular markers? *Theoretical and Applied Genetics*, 111(7), 1288–1299. <https://doi.org/10.1007/s00122-005-0055-7>
- Lanes, E. C. M., Viana, J. M. S., Paes, G. P., Paula, M. F. B., Maia, C., Caixeta, E. T., & Miranda, G. V. (2014). Population structure and genetic diversity of maize inbreds derived from tropical hybrids. *Genetics and Molecular Research*, 13(3), 7365–7376. <https://doi.org/10.4238/2014.September.12.2>
- Leff, B., Ramankutty, N., & Foley, J. A. (2004). Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, 18(1), n/a-n/a. <https://doi.org/10.1029/2003gb002108>
- Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, J. G., Buckler, E., & Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9), 6080–6084. <https://doi.org/10.1073/pnas.052125199>
- Mir, C., Zerjal, T., Combes, V., Dumas, F., Madur, D., Bedoya, C., Dreisigacker, S., Franco, J., Grudloyma, P., Hao, P. X., Hearne, S., Jampatong, C., Laloë, D., Muthamia, Z., Nguyen, T., Prasanna, B. M., Taba, S., Xie, C. X., Yunus, M., ... Charcosset, A. (2013). Out of America: tracing the genetic footprints of the global diffusion of maize. *Theoretical and Applied Genetics*, 126(11), 2671–2682. <https://doi.org/10.1007/s00122-013-2164-z>

- Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the New York Academy of Sciences*, 1312(1), 105–112. <https://doi.org/10.1111/nyas.12396>
- Serratos-Hernández, J. A. (2009). The origin and diversity of maize in the american continent. <https://doi.org/10.2166/wst.2009.007>
- Stevenson, J. C., & Goodman, M. M. (1972). Ecology of Exotic Races of Maize. I. Leaf Number and Tilling of 16 Races Under Four Temperatures and Two Photoperiods 1 . *Crop Science*, 12(6), 864–868. <https://doi.org/10.2135/cropsci1972.0011183x001200060045x>
- Szabó, V. M., & Burr, B. (1996). Simple inheritance of key traits distinguishing maize and teosinte. *Molecular and General Genetics*, 252(1–2), 33–41. <https://doi.org/10.1007/BF02173202>
- Wu, Y., San Vicente, F., Huang, K., Dhliwayo, T., Costich, D. E., Semagn, K., Sudha, N., Olsen, M., Prasanna, B. M., Zhang, X., & Babu, R. (2016). Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theoretical and Applied Genetics*, 129(4), 753–765. <https://doi.org/10.1007/s00122-016-2664-8>