GENOME-WIDE ASSOCIATION STUDY FOR BACTERIAL LEAF STREAK RESISTANCE IN MAIZE

Ruiz Marcos[1], Rossi Ezequiel A.[1], Bonamico Natalia C.[1], Balzarini Mónica G.[2]

[1] Mejoramiento Genético, FAV, Universidad Nacional de Río Cuarto, Argentina, Instituto de Investigaciones Agrobiotecnológicas (INIAB, CONICET-UNRC)

[2] Estadística y Biometría, FCA, Universidad Nacional de Córdoba, Argentina, Unidad de Fitopatología y Modelización Agrícola (UFYMA, CONICET-INTA)

Email: mruiz@ayv,unrc.edu.ar

Abbreviations:

BLS, Bacterial leaf streak; CIMMYT, International Maize and Wheat Improvement Center, GWAS, genome wide association studies; LD, linkage disequilibrium; GLM, general linear model; MLM, mixed linear model; SEV, severity; SNP: single nucleotide polymorphism; MVA: mevalonic acid.

Core ideas

GWAS detected genomic regions linked to bacterial leaf streak severity in maize.

Each region explained 10 to 17% of the observed severity variability.

The identified alleles are useful to improve maize breeding programs.

ABSTRACT

Maize (*Zea Mays* L), one of the most important crops worldwide, is affected by foliar diseases that limit global production. Bacterial diseases have increased in Argentina during the last years. The aim of this work was to explore a maize panel provided by the

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/agj2.21299.

International Maize and Wheat Improvement Center (CIMMYT) to identify alleles associated with resistance to bacterial leaf streak of maize (BLS), a disease caused by *Xanthomonas vasicola* pv *vasculorum,* in central Argentina. A diverse panel of 200 maize lines was evaluated for resistance to bacterial diseases in four environments of central Argentina in 2020 and 2021. The predictor of the genetic merit that does not include environmental effects and 46,990 SNPs obtained by genotyping-by-sequencing were used in the genome-wide association study (GWAS). The ten lines with the lowest severity across environments belonged to different environmental adaptation programs defined by CIMMYT. The GWAS allowed us to identify eleven genomic regions associated with BLS, located in chromosomes 1, 2, 5, 7, 8 and 9. Five of those regions, located in bins 1.04, 2.01, 5.03, 8.06 and 9.03 were associated with plant defense candidate genes such as, strictosidine synthase-like 11, protein-serine/threonine phosphatase and a putative LRR receptor-like serine/threonine-protein kinase gene. Our study provides potential resistance alleles to BLS that can be incorporated to improve maize breeding programs.

1. INTRODUCTION

Maize (*Zea Mays* L) is one of the most important staple food and animal feed crops worldwide (Mideros et al., 2014). Argentina is the fourth largest producer, with an annual production of 58.3 million tons (FAOSTAT, 2022). The crop is affected by several pathogens that cause different diseases and reduce production worldwide (Yang et al., 2017). Foliar diseases are a significant limiting factor in maize production globally (Pratt and Gordon, 2006). Since 2010, bacterial leaf streak of maize (BLS), caused by *Xanthomonas vasicola pv. vasculorum,* has expanded its geographic range in the central region of Argentina. Symptoms are initially expressed as water-soaked lesions on the leaves that expand along the veins, producing irregular long necrotic streaks and dark yellow to brown lesions (Plazas et al.,

2018). The most economical and environmentally acceptable means of reducing losses due to maize diseases is the use of genotypes with genetic resistance (Carson et al., 2004). Thus, studies are increasingly addressing the genetic basis of maize reaction to viral, fungal, and bacterial diseases. However, the genetics of maize response to bacterial infection has received less attention than that of fungal and viral diseases (Rossi et al., 2019).

Genome-wide association studies (GWAS) have proved to be a powerful tool to identify specific allele variants that confer resistance to diseases (Zila et al., 2014). This approach takes full advantage of the natural variation within a germplasm collection to identify the genetic loci underlying traits at a relatively high resolution (Chen et al., 2015; Yu and Buckler, 2006). GWAS are based on the analysis of the statistical association between genotypic marker alleles determined in a group of genotypes and the phenotype under study. Maize is an ideal crop for GWAS because it has abundant genetic diversity and rapid linkage disequilibrium (LD) decay (Chen et al., 2015). There are some particular concerns related to population structure, which can cause allele frequencies to differ significantly between subpopulations, creating unexpected LD between unlinked loci across the genome (Guo et al., 2014). When population structure is not accounted for, spurious associations may be detected between disease resistance and alleles at other loci that are all differentially distributed among subpopulations. Moreover, GWAS require dense genotyping with genetic marker loci covering all of the chromosomes (Rafalski, 2010) to increase the odds to identify linked alleles as well as to quantify the underlying genetic population structure.

The maize lines provided by the International Maize and Wheat Improvement Center (CIMMYT) maize lines are one of the most widely distributed genetic resources of publicly generated elite lines, which are freely available to both public and private sector breeders (Chen et al., 2016). These lines have been molecularly characterized with SNPs obtained using a Genotyping-by-Sequencing approach. A highly diverse panel of these lines was

evaluated in central Argentina to identify alleles that may be useful to enhance maize breeding for resistance to endemic viral and fungal diseases (Rossi, Ruiz et al., 2020). In this work, the CIMMYT panel, whose material is exotic to Argentina, was explored with the aim to identify alleles associated with resistance to bacterial disease caused by *Xanthomonas vasicola* pv. *vasculorum* in central Argentina.

2. MATERIALS AND METHODS

2.1. Plant material and field trial

A diverse panel of 200 maize lines from CIMMYT was evaluated for resistance to BLS in four environments (Río Cuarto 2020, Río Cuarto 2021, Sampacho 2021, and La Cautiva 2021) of central Argentina in 2020 and 2021. These inbred lines represent the three major environmental adaptations from the CIMMYT germplasm collection (Wu et al., 2016). The trials were conducted under natural infection. In each environment, the susceptible maize line L420 was planted in each block as susceptible control. In Río Cuarto 2020 plots were 2.5 m long, with a row spacing of 0.52 m, while in the remaining environments plots were 5 m long with a row spacing of 0.52 m. A partially replicated (p-rep) design was used (Cullis et al., 2006), with 25% of the genotypes with three replications and the remaining genotypes with one replication. All plots were hand-weeded as needed.

2.2. Phenotypic data

All plants in each plot were evaluated and scored for BLS by the presence of irregular long necrotic streaks, and dark yellow to brown lesions at flowering stage. Each plant was classified by the degree of disease severity, according to the ordinal scale proposed by Schuelter et al. (2003): 1 = no symptoms; 2 = scattered lesions; 3 = up to 50% of the lower leaves presenting lesions, with severe injuries in 25% of the lower leaves; 4 = up to 75% of the leaves presenting lesions, with severe injuries in 50% of the lower leaves; 5 = 100% of

the leaves with lesions, with severe injuries in 75% of the lower leaves. The incidence was estimated as the proportion of diseased plants in each plot. The severity (SEV) per plot was estimated as the mean rating of all plants in the plot.

2.3. Genomic data

The genotypic characterization used for this study was performed using SNP markers (Wu et al., 2016) available from http://data.cimmyt.org/dvn. Of a total of 362,008 SNPs, minor SNP states, SNPs with minor allele frequency less than 0.15 and with high missing data rate (>35%) were removed. Only 46,990 SNPs which were distributed in the 10 chromosomes were kept.

2.4. Statistical analysis

2.4.1. Phenotypic data analysis

Analysis of variance for SEV was performed by integrating data across environments in a mixed linear model with the R/sommer package (Covarrubias-Pazaran, 2016). The model included environment, genotype, block within environment, and the genotype-by-environment interaction effects. Environment effect was regarded as fixed, and the remaining factors were fit as random effects.

$$y_{ijk} = \mu + G_i + E_j + b_{kj} + GE_{ij} + e_{ijk}$$

where y_{ijk} is the vector of phenotypic data, μ is the overall mean, G_i is the random effect of genotype i, E_j is the fixed effect of the environment j, b_{kj} is the nested effect of the block k within environment j, GE_{ij} is the random effect of the interaction between genotype i and environment j, e_{ijk} is the random effect of error, assuming that G_{ijk} ~ ~ , and e_{ijk} ~).

The variance components (REML estimates) were used to calculate mean-basis heritability, as proposed by Hallauer and Miranda (1988).

In each environment:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_e^2}{p}\right)}$$

Across environments:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2\left(\frac{\sigma_{ge}^2}{e}\right) + \left(\frac{\sigma_e^2}{p}\right)}$$

where σ_g^2 is the genotypic variance, σ_{ge}^2 is the variance of interaction between genotype and environment, σ_e^2 is the error variance, e is the number of environments, and p is a weighted mean of the number of replications per genotype in each environment and across environments. The weighting used was that proposed by Holland et al. (2010). The mixed model was also used to estimate the best linear unbiased predictor (BLUP) of genotype effects, i.e., a predictor of the genetic merit that does not include environmental effects. The BLUPs of genotypic effects on SEV were used in the GWAS as response variables.

2.4.2. Genome-wide association study

Association tests for each SNP were performed using the software TASSEL 5.2.59 (Bradbury et al., 2007). For GWAS, alternative models were assessed: 1) a general linear model (GLM) (Naïve model; 2) a mixed linear model (MLM) including the kinship matrix (Parisseaux and Bernardo, 2004) to model genetic relationship between any two lines in the studied population (K model); 3) a GLM including five principal components of genomic data as a strategy to model the underlying genetic structure in the population (PCA model) (Zhao et al., 2007); 4) a GLM using the Q matrix obtained from the software STRUCTURE (Pritchard et al., 2000) to account for the genetic population structure (Q model); 5) a MLM including both the genetic structure suggested by PCA and the kinship matrix (PCA + K model) (Yu et al., 2006); and 6) a MLM including both the genetic structure denoted by STRUCTURE and the kinship matrix (Q + K model). All fittings were compared, and the

most suitable model was selected by quantile-quantile plots, which compare the observed -log10 (P-value) with the expected -log10 (P-value) under the null-hypothesis of no associations between marker and the studied trait. The SNP markers were determined to be significantly associated with SEV at a threshold of -log10 (P-value) > 3.5 (P-value < 0.0003165) estimated using the number of independent tests in the marker data as proposed by Li and Ji (2005). Manhattan and quantile–quantile plots were created by qqman package (Turner, 2018) in R software using the GWAS results. A multiple linear regression analysis was performed with the associated markers to estimate the proportion of total phenotypic variance explained by the identified SNPs.

3. RESULTS

In the four environments evaluated, water-soaked lesions producing irregular long necrotic streaks and dark yellow to brown lesions were observed in maize lines (Fig.1). These symptoms were compatible with those caused by *Xanthomonas vasicola* pv. *vasculorum*. Control susceptible maize line showed 100% of incidence and the highest degree of disease severity per plot in all environments. Incidence of BLS, measured as the percentage of plants showing symptoms of BLS within each plot, ranged from 83% to 96% in all environments, with a mean incidence of 91% across environments. The average plot SEV was normally distributed (Fig.1 and Table 1). The Río Cuarto 2020 environment had the lowest average plot SEV (2.52), whereas La Cautiva 2021 environment had the highest average plot SEV (3.46). Heritability was moderate to high within environments and across environments, with a range of 0.43 to 0.73 (Table 1).

The 5% of lines with the lowest SEV across environments belong to the Subtropical, Lowland, Asia lowland, Africa lowland or Africa MA/ST environmental adaptation (Table 2 and Supplemental table 1). While, the 5% of lines with the highest SEV belong to the

Lowland, Subtropical, Africa MA/ST and South América environmental adaptation (Supplemental table 1).

As shown in the quantile-quantile plot, six linear models were tested, and the most suitable for our data was the one using kinship relationship (K matrix) and principal components analysis (PCA) as covariate to account for the underlying population genetic structure. A total of 15 significantly associated SNPs for BLS were detected using the PCA+K model (Figure 2).

The significant SNPs within a 20-kb region were considered to represent a locus, and the SNP with the lowest P-value in a locus was defined as the SNP with the closest linkage to the causal gene. According to this, 11 associations were detected (Table 3). The significant SNPs were located in chromosome 1 (two SNPs located in bins 1.04 and 1.11), chromosome 2 (two SNPs located in bins 2.01 and 2.04), chromosome 5 (three SNPs located in bins 5.03), chromosome 7 (one SNP located in bin 7.01), chromosome 8 (two SNPs located in bins 8.03 and 8.06), and chromosome 9 (one SNP located in bin 9.03. The significantly associated SNPs individually explained 10 to 17% of the total phenotypic variance (Table 3). A multiple linear regression analysis with the significantly associated markers and the adjusted means explained 50% of the phenotypic variation of SEV.

DISCUSSION

The most cost-effective and environmentally acceptable means of reducing losses due to maize diseases is the use of genotypes with genetic resistance (Carson et al., 2004). CIMMYT maize lines showed great phenotypic and genotypic variability for resistance to diseases caused by bacterial agents in Argentina. The heritability values obtained in this work ranged between 0.43 and 0.73. When compared with heritability values obtained by Qiu, Kaiser et al. (2020), who studied BLS in three nested association mapping populations and their parental lines, our heritability values were similar in La Cautiva 2021 and higher in the

remaining environments and across environments The values for the estimated mean-basis heritability reveal predominance of additive control in responses from the maize inbred lines to BLS and favor the power of quantitative trait loci (QTL) detection, as suggested by Yu et al. (2008).

In this work, we identified 11 SNPs significantly linked to BLS resistance. Genomic regions for BLS resistance were located in bins 1.04, 1.11, 2.01, 2.04, 5.03, 7.01, 8.03, 8.06, and 9.03. By linkage mapping, Qiu, Kaiser et al. (2020) found five significant QTL for BLS located in bins 1.05, 2.03, 3.08, 4.07 and 5.05. Even though we evaluated the same disease, our findings did not match those of these authors. This might be due to the fact that their resistant donor parents are not included in our diverse panel. Also, as stated by Qiu, Kaiser et al. (2020), the lack of common QTL is due to the low marker coverage in one of their populations that may have impeded QTL identification, and that populations were evaluated in different environments using different bacterial strains of the same disease.

Of the identified SNPs, seven genomic regions were reported by different authors for resistance to bacterial diseases (Brown et al., 2001; Gomes de Paula Lana et al., 2017; Cooper et al., 2018; Rossi, Kuki et al., 2020, Qiu, Cooper et al., 2020). Our results agree with previous works; indeed, Brown et al. (2001) identified a genomic region in bin 5.03 associated with resistance to Stewart´s wilt; Cooper et al. (2018) identified a genomic region associated with resistance to Goss's wilt in bin 7.01; Gomes de Paula Lana et al. (2017) identified a genomic region associated with resistance to maize white spot in bin 8.03; and Rossi, Kuki et al. (2020) identified a genomic region associated with maize white spot resistance in bin 8.06; and Qiu, Cooper et al. (2020) identified a genomic region associated with resistance to Goss's wilt in bin 9.03.

Here, we identified genes that may be involved in plant disease. In bin 1.04, we found a genomic region adjacent to Strictosidine synthase-like 11. Strictosidine is a precursor in

many pathways in different plants that leads to the production of about 2000 alkaloid compounds (Hicks et al., 2011). In *Catharanthus roseus*, Luijendijk et al. (1996) showed that the deglucosilation of strictosidine forms an antimicrobial compound that reduces bacterial growth. In bin 2.01, we found a gene with protein-serine/threonine phosphatase predicted function. Serine-Threonine Phosphatases are a wide family of enzymes with important roles in signal transduction pathways, the regulation of cell cycle and metabolism, stress responses and defense (Máthé et al., 2019). Additionally, in bin 5.03, a putative LRR receptor-like serine/threonine-protein kinase gene, was identified. Many functions have been attributed to these receptors, from plant growth and development to symbiosis and immunity (Dievart et al., 2020). Moreover, in bin 8.06, we identified a putative cytochrome P450 superfamily protein. This superfamily protein plays an important role in plant defense, since it is involved in phytoalexin biosynthesis, hormone metabolism and the biosynthesis of some other secondary metabolites (Xu et al., 2015). Furthermore, in bin 9.03 we identified a putative E3 ubiquitin ligase. In Arabidopsis, E3 ubiquitin ligase is a positive regulator of 3-Hydroxy-3-methylglutaryl coenzyme A reductase activity. This enzyme controls the metabolic flux in the early steps of the mevalonic acid (MVA) pathway (Doblas et al., 2013). In plants, the MVA pathway is involved in isoprenoid biosynthesis that play a vital role in plant defense responses and participate in fundamental physiological and developmental processes (Hemmerlin et al., 2012). The reported genomic regions with unknown function are promising for further analysis.

The 11 genomic regions identified in this study may contain alleles that can be incorporated into maize breeding programs. Further research is required to validate these genomic regions. The stacking of favorable alleles using marker-assisted backcross breeding should be a useful approach to incorporate resistance in susceptible germplasm.

SUPPLEMENTAL MATERIAL

Table S 1. CIMMYT maize lines (CML) adaptation program, grain colour, grain texture and average plot severity (SEV) of bacterial disease assessed in maize plants across environments of central Argentina.

ACKNOWLEDGMENTS

We would like to acknowledge the International Maize and Wheat Improvement Center (CIMMYT) who provided us their maize lines for research.

FUNDING

This work was supported by the Agencia Nacional de Promoción Científica y Tecnológica, PICT-03321/2018; the Secretaría de Ciencia y Técnica, Universidad Nacional de Río Cuarto PPI 083/20 and the National Scientific and Technical Research Council (CONICET).

AUTHOR CONTRIBUTION STATEMENT

EAR and NCB requested the plant material from CIMMYT; MR, EAR, NCB, and MB designed the experiment; MR, EAR, and NCB conducted the field evaluations; MR and EAR carried out the statistical analysis of phenotypic data and GWAS analysis under MB supervision; MR, EAR, NCB, and MB interpreted the results and drafted the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

4. REFERENCES

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*, 2633– 2635. https://doi.org/10.1093/bioinformatics/btm308

Brown, A. F., Juvik, J. A., & Pataky, J. K. (2001). Quantitative trait loci in sweet corn associated with partial resistance to Stewart's wilt, northern corn leaf blight, and common rust. *Phytopathology, 91(3)*, 293–300. https://doi.org/10.1094/PHYTO.2001.91.3.293

Carson, M. L., Stuber, C. W., & Senior, M. L. (2004). Identification and mapping of quantitative trait loci conditioning resistance to Southern leaf blight of maize. *Phytopathology 94(8)*, 862-867.

Chen, G., Wang, X., Hao, J., Yan, J., & Ding, J. (2015). Genome-wide association implicates candidate genes conferring resistance to maize rough dwarf disease in maize. *PLoS ONE, 10(11)*, 1–13. https://doi. org/10.1371/journal.pone.0142001

Chen, J., Zavala, C., Ortega, N., Petroli, C., Franco, J., Burgueño, J., Costich, D. E., & Hearne, S. J. (2016). The development of quality control genotyping approaches: A case study using elite maize lines. *PLoS ONE, 11(6)*. https://doi.org/10.1371/journal.pone.0157236

Cooper, J. S., Balint-Kurti, P. J., & Jamann, T. M. (2018). Identification of quantitative trait loci for Goss's wilt of maize. *Crop Science, 58(3)*, 1192–1200. https://doi.org/10.2135/cropsci2017.10.0618

Covarrubias-Pazaran, G. (2016). Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS ONE, 11*, e0156744. https://doi.org/10.1371/journal.pone.0156744

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics, 11(4)*, 381–393. https://doi.org/10.1198/108571106X154443

Dievart, A., Gottin, C., Périn, C., Ranwez, V., & Chantret, N. (2020). Origin and diversity of plant receptor-like kinases. *Annual Review of Plant Biology, 71,* 131-156 https://doi.org/10.1146/annurev-arplant-073019

Doblas, V. G., Amorim-Silva, V., Posé, D., Rosado, A., Esteban, A., Arró, M., Azevedo, H., Bombarely, A., Borsani, O., Valpuesta, V., Ferrer, A., Tavares, R. M., & Botellaa, M. A. (2013). The SUD1 gene encodes a putative E3 ubiquitin ligase and is a positive regulator of 3-hydroxy-3-methylglutaryl coenzyme a reductase activity in Arabidopsis. *Plant Cell, 25(2)*, 728–743.

FAOSTAT. (2022). Food and Agricultural Organization of the United Nations. Maize Available in http://www.fao.org/faostat/en/#data/QC (consulted January 10. 2022).

Gomes de Paula Lana U., Prazeres de Souza I. R., Noda R. W., Pastina M. M., Vieira Magalhaes J., & Teixeira Guimaraes C. (2017). Quantitative trait loci and resistance gene analogs associated with maize white spot resistance. *Plant Disease, 101*, 200-208. doi:10.1094/PDIS06-16-0899-RE.

Guo, Z., Tucker, D. M., Basten, C. J., Gandhi, H., Ersoz, E., Guo, B., Xu, Z., Wang, D., & Gay, G. (2014). The impact of population structure on genomic prediction in stratified populations. *Theoretical and Applied Genetics, 127(3),* 749–762. https://doi.org/10.1007/s00122-013-2255-x

Hallauer, A. R., & Miranda, J. B. (1988). Quantitative genetics in maize breeding (2nd Ed.). Ames, IA: Iowa State University Press.

Hemmerlin, A., Harwood, J. L., & Bach, T. J. (2012). A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? *Progress in Lipid Research, 51(2)*, 95–148. https://doi.org/10.1016/j.plipres.2011.12.001

Hicks, M. A., Barber, A. E., Giddings, L. A., Caldwell, J., O'Connor, S. E., & Babbitt, P. C. (2011). The evolution of function in strictosidine synthase-like proteins. *Proteins, 79(11),* 3082–3098. https://doi.org/10.1002/prot.23135

Holland, J. B., Nyquist, W. E., & Cervantes-Martínez, C. T. (2010). Estimating and interpreting heritability for plant breeding: An update. In J. Janick (Ed.), Plant Breeding

Reviews. Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470650202.ch2

Li J., & Ji L. (2005). Adjusting multiple testing in multilocus analyses using the eigen values of a correlation matrix. *Heredity 95,* 221-227. https://doi.org/10.1038/ sj.hdy.6800717

Luijendijk, T. J. C., van der Meijden, E. D., & Verpoorte, R. (1996). Involvement of strictosidine as a defensive chemical in *Catharanthus roseus*. *Journal of Chemical Ecology, 22 (8).*

Máthé, C., Garda, T., Freytag, C., & Hamvas, M. M. (2019). The role of serine-threonine protein phosphatase pp2a in plant oxidative stress signaling—facts and hypotheses. *International Journal of Molecular Sciences, 20, (12).* https://doi.org/10.3390/ijms20123028

Mideros, S. X., Warburton, M. L., Jamann, T. M., Windham, G. L., Paul Williams, W., & Nelson, R. J. (2014). Quantitative trait loci influencing mycotoxin contamination of maize: Analysis by linkage mapping, characterization of near-isogenic lines, and meta-analysis. *Crop Science, 54(1),* 127–142. https://doi.org/10.2135/cropsci2013.04.0249

Parisseaux B., & Bernardo R. (2004) In silico mapping of quantitative trait loci in maize. *Theoretical and Applied Genetics, 109*, 508- 514

Plazas, M.C., De Rossi R.L., Brücher E., Guerra F.A., Vilaró M., Guerra G.D., Wu G., Ortiz-Castro M.C., & Broders K. (2018). First report of *Xanthomonas vasicola* pv. vasculorum causing bacteria leaf streak of maize (Zea mays) in Argentina. *Plant Disease, 102(5),* 1026. https://doi.org/10.1094/PDIS-10-17-1578-PDN

Pratt, R. C., & Gordon, S. G. (2006). Breeding for resistance to maize foliar pathogens. In J. Janick (Ed.), Plant breeding reviews (pp. 119– 173). Oxford: John Wiley & Sons.

Pritchard J., Stephens M., Rosenberg N., & Donnelly P. (2000). Association mapping in structured populations. *American Journal of Human Genetics, 67*, 170-181.

Qiu, Y., Cooper, J., Kaiser, C., Wisser, R., Mideros, S. X., Jamann, T. M. (2020). Identification of loci that confer resistance to bacterial and fungal diseases of maize. *G3 Genes, genomes, genetics 10,* 2819-2228.

Qiu, Y., Kaiser, C., Schmidt, C., Broders, K., Robertson, A. E., S. X., Jamann, T. M. (2020). Identification of quantitative trait loci associated with maize resistance to bacterial leaf streak. *Crop Science, 60,* 226-237.

Rafalski, J. A. (2010). Association genetics in crop improvement. *Current Opinion in Plant Biology, 13(2)*, 174–180. https://doi.org/10.1016/j.pbi.2009.12.004

Rossi, E. S., Kuki, M. C., Pinto, R. J. B., Scapim, C. A., Faria, M. V., & de Leon, N. (2020). Genomic-wide association study for white spot resistance in a tropical maize germplasm. *Euphytica, 216(1).* https://doi.org/10.1007/s10681-019-2550-y

Rossi, E. A., Ruiz, M., Bonamico, N. C., & Balzarini, M. G. (2020). Identifying inbred lines with resistance to endemic diseases in exotic maize germplasm. *Crop Science. 60*, 3141-3150. https://doi.org/10.1002/csc2.20275

Rossi, E. A., Ruiz, M., Rueda Calderón, M. A., Bruno, C. I., Bonamico, N. C., & Balzarini, M. G. (2019). Meta-analysis of QTL studies for resistance to fungi and viruses in maize. *Crop Science, 59(1)*, 125–139. https://doi.org/10.2135/cropsci2018.05.0 330

Schuelter, A. R., Regina, I., Souza, P. D. E., Tavares, F., Xavier, M., Santos, D. O. S., Oliveira, E. D. E., & Guimarães, C. T. (2003). Genetic control of maize resistance to Phaeosphaeria. *Revista Brasileira de Milho e Sorgo, 2(1)*, 80–86.

Turner S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *Journal of Open Source Software*

Wu, Y., San, F., Kaijian, V., & Thanda, H. (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

Xu, J., Wang, X. Y., & Guo, W. Z. (2015). The cytochrome P450 superfamily: Key players in plant development and defense. *Journal of Integrative Agriculture, 14(9)*, 1673–1686. https://doi.org/10.1016/S2095-3119(14)60980-1

Yang, Q., Balint-Kurti, P., & Xu, M. (2017). Quantitative disease resistance: Dissection and adoption in maize. *Molecular Plant, 10(3),* 402– 413. https://doi.org/10.1016/j.molp.2017.02.004

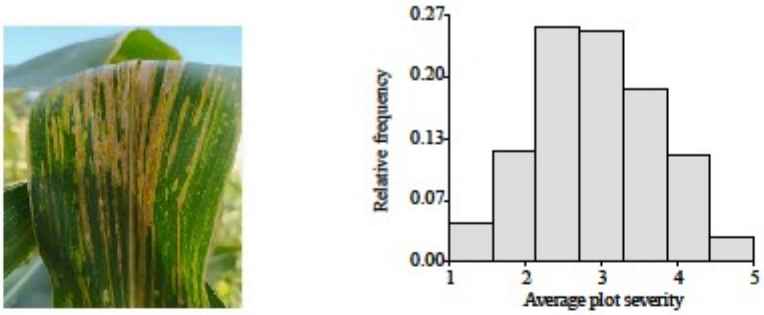
Yu, J., & Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology, 17(2)*, 155–160. https://doi.org/10.1016/j.copbio.2006.02.003

Yu, J., Holland, J. B., McMullen, M. D., & Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics, 178*, 539–551. https://doi.org/10.1534/ genetics.107.074245

Yu J., Pressoir G., Briggs W., Bi I., Yamasaki M., Doebley J., McMullen M., Gaut B., Holland J., Kresovich S., & Buckler E. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics 38*, 203-208.

Zhao K., Aranzana M., Kim S., Lister C., Shindo C., Tang C., Toomajian C., Zheng H., Dean C., Marjoram P., & Nordborg M. (2007). An Arabidopsis example of association mapping in structured samples. *PLoS Genetics 3*, e4.

Zila, C. T., Ogut, F., Romay, M. C., Gardner, C. A., Buckler, E. S., & Holland, J. B. (2014). Genome-wide association study of Fusarium ear rot disease in the U.S.A. maize inbred line collection. *BMC Plant Biology, 14(1)*. https://doi.org/10.1186/s12870-014-0372-6



a) b)

Figure 1. Symptoms observed in maize leaves (a) and frequency distribution of average plot severity (b) of bacterial disease assessed in maize plants across environments of central Argentina.

Figure 1. Symptoms observed in maize leaves (a) and frequency distribution of average plot severity (b) of bacterial disease assessed in maize plants across environments of central Argentina.

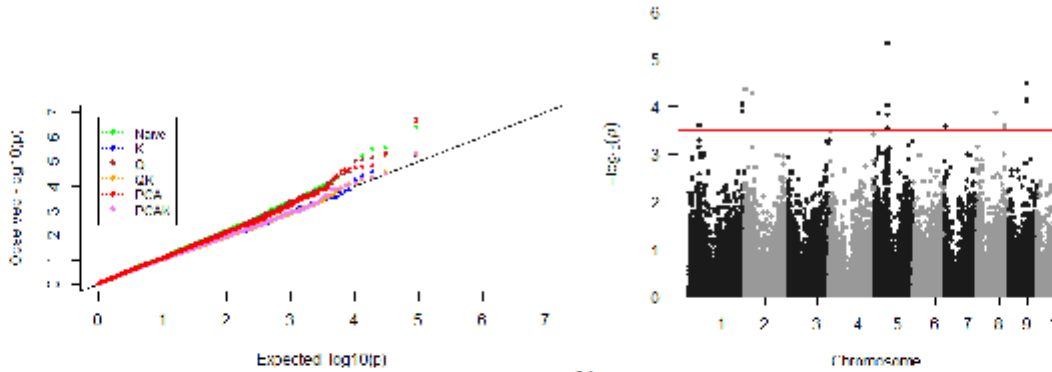


Figure 2. Quantile-quantile plot (a) and Manhattan plot (b) of GWAS results for severity of bacterial leaf streak disease in m
in central Argentina. The black line in the quantile-quantile plot indicates the expected p-value distribution under the null hypoth
markers. The horizontal line in the Manhattan plot depicts the significance threshold (P = 0.000316). The x-axis indicates the S
the 10 chromosomes.

Figure 2. Quantile-quantile plot (a) and Manhattan plot (b) of GWAS results for severity of bacterial leaf streak disease in maize lines assessed in central Argentina. The black line in the quantile-quantile plot indicates the expected p-value distribution under the null hypothesis of no causative markers. The horizontal line in the Manhattan plot depicts the significance threshold (P = 0.000316). The x-axis indicates the SNP position along the 10 chromosomes.

TABLES

Table 1. Mean and heritability (H^2) of maize inbred lines evaluated for severity bacterial leaf streak of maize in four environments and across environments of central Argentina.

Environment	mean ± S.E.	H^2
Río Cuarto 2020	2.52 ± 0.06	0.65
Río Cuarto 2021	2.93 ± 0.06	0.73
Sampacho 2021	3.10 ± 0.05	0.59
La Cautiva 2021	3.46 ± 0.06	0.43
Across environments	2.97 ± 0.05	0.53

Table 2. Selected lines with lower severity of bacterial leaf streak of maize across environments of central Argentina.

Genotype	Environmental adaptation	Grain colour	Grain texture	SEV[1]
CML_94	Subtropical	White	Semident	1.00
CML_96	Subtropical	White	Dent	1.25
CML_65	Lowland	White	Semident	1.55
CML_433	Asia Lowland	Yellow	Flint	1.56
CML_228	Africa Lowland	Yellow	Flint	1.56
CML_249	Lowland	White	Dent	1.57
CML_224	Africa Lowland	Yellow	Flint	1.60
CML_389	Africa MA/ST	White	Flint	1.62
CML_388	Africa MA/ST	White	Flint	1.66
CML_484	Subtropical	White	Semident	1.70

[1]SEV: severity of bacterial disease assessed in maize plants across environments of central Argentina.

Table 3. SNP markers associated with severity of maize bacterial disease in central Argentina.

Marker	Bin	Allele	Additive effect	P-value	R^2	Gene model	Predicted fun
S1_58459485	1.04	C/G	$2x10^{-5}$	$2.49x10^{-04}$	0.14	Zm00001d029115	Strictosidine syntha
S1_289580805	1.11	G/A	0.18	$1.29x10^{-04}$	0.17	Zm00001d034330	Uncharacterized
S2_3990225	2.01	T/C	-0.09	$4.32\ x10^{-05}$	0.15	Zm00001d001988	Protein-serine/th phosphatase
S2_45194157	2.04	G/A	0.06	$5.30x10^{-05}$	0.12	Zm00001d003464	DUF4057 domain- protein
S5_29622018	5.03	A/T	0.26	$1.37x10^{-04}$	0.15	Zm00001d014032	Protein light-depen hypocotyls 3
S5_76381650	5.03	A/G	-0.24	$9.23x10^{-05}$	0.10	Zm00001d015120	RuvB-like hel
S5_77995036	5.03	A/G	0.21	$4.66x10^{-06}$	0.10	Zm00001d015163	Putative LRR rece serine/threonine-protein
S7_15093203	7.01	C/T	0.26	$2.55x10^{-04}$	0.10	Zm00001d018785	Serrate RNA effecto
S8_104660128	8.03	T/C	0.16	$1.34x10^{-04}$	0.13	Zm00001d010225	DEAD-box ATP-dep helicase 17
S8_155081879	8.06	T/C	0.20	$2.58\ x10^{-04}$	0.10	Zm00001d011586	Putative cytochro superfamily pro
S9_99365447	9.03	C/G	0.03	$7.41\ x10^{-05}$	0.10	Zm00001d046624	Putative E3 ubiquitin