KARLA JORGE DA SILVA

DIVERSIDADE GENÉTICA E MAPEAMENTO ASSOCIATIVO PARA RESISTÊNCIA À FUMONISINAS EM MILHO TROPICAL UTILIZANDO MARCADORES SNPs

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

Orientador: Luiz Antonio dos Santos Dias

Coorientador: Lauro José Moreira Guimarães

VIÇOSA - MINAS GERAIS

2019

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T
S586d
2019

Silva, Karla Jorge da, 1990-
Diversidade genética e mapeamento associativo para resistência à fumonisinas em milho tropical utilizando marcadores SNPs / Karla Jorge da Silva. – Viçosa, MG, 2019.
95f. : il. (algumas color.) ; 29 cm.

Orientador: Luiz Antonio dos Santos Dias.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Melhoramento genético. 2. Marcadores moleculares. 3. Genes. 4. *Zea mays*. I. Universidade Federal de Viçosa. Departamento de Fitotecnia. Doutorado em Genética e Melhoramento. II. Título.

CDD 22 ed. 633.152

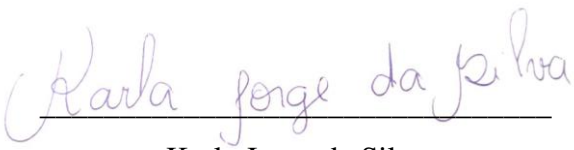
KARLA JORGE DA SILVA

DIVERSIDADE GENÉTICA E MAPEAMENTO ASSOCIATIVO PARA RESISTÊNCIA À FUMONISINAS EM MILHO TROPICAL UTILIZANDO MARCADORES SNPs

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 20 de dezembro 2019.

Assentimento:



Karla Jorge da Silva
Autora

Luiz Antonio dos Santos Dias
Orientador

A Deus,
Aos meus pais, Carlos (in memoriam) e Eliane,
À minha irmã Michele
Ao meu esposo, Daniel.
Dedico

AGRADECIMENTOS

Agradeço primeiramente a DEUS, pela vida e por proporcionar tudo que sou e tenho.

Aos meus pais, Eliane Jorge da Silva e Carlos Roberto Jorge, e a minha irmã Michele Jorge pelo apoio, incentivo e por todo amor e carinho a mim dedicados.

Ao meu esposo, Daniel Pereira de Paula, por todo o amor, carinho, companheirismo e ajuda ao longo da minha formação acadêmica, sendo o meu suporte durante esta caminhada.

Ao meu orientador Prof. Luiz Antônio Dias pelo brilhantismo, humildade e por ser um exemplo de profissional.

A Pesquisadora Maria Marta Partina pelos ensinamentos, amizade e por ser um modelo de profissional a ser seguido.

Ao Pesquisador Dr. Lauro José Moreira Guimarães, pela orientação na Embrapa Milho e Sorgo e conhecimentos transmitidos.

A Pesquisadora Claudia Teixeira Guimarães pela aprendizagem, disponibilidade e suporte no decorrer do doutorado.

Ao Roberto Trindade e Sylvia Moraes pelos ensinamentos e apoio.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento pela oportunidade e excelência em ensino.

À Embrapa Milho e Sorgo, por possibilitar que os experimentos fossem conduzidos na unidade.

Ao CNPq, pela concessão da bolsa de estudos.

Ao amigos do laboratório de Bioinformática, em especial Karine e Jose Henrique, pela amizade, pelos ensinamentos e disponibilidade para me atender quando as dúvidas surgiram sempre me trazendo calma.

Ao laboratório de Segurança Alimentar e Fitopatologia da Embrapa Milho e Sorgo.

À equipe do Galpão de Melhoramento de Milho, pelo apoio na condução dos experimentos.

A todos os professores, funcionários e amigos da UFV pelo aprendizado e pela convivência.

Enfim, todos os familiares e amigos, que direta ou indiretamente contribuíram para que eu chegasse até aqui, e pudesse realizar esse grande sonho, o meu sorriso e o meu carinho. Muito obrigada!

“The best of all things is to learn.
Money can be lost or stolen, health and strength may fail but what you have committed to your mind is yours forever.”
(Louis L’Amour)

BIOGRAFIA

Karla Jorge da Silva, filha de Carlos Roberto Jorge (*in memoriam*) e Eliane Jorge da Silva, nasceu na cidade de Matozinhos, Minas Gerais, Brasil, em 16 de maio de 1990.

Em fevereiro de 2014 obteve o título de Engenheira Agrônoma pela Universidade Federal de São João Del-Rei, em Sete Lagoas, Minas Gerais, Brasil.

Em março do mesmo ano ingressou no curso de Mestrado no Programa de Pós Graduação em Genética e Melhoramento, na Universidade Federal de Viçosa e concluiu em 2016.

Em março do ano de 2016 ingressou no Doutorado em Genética e Melhoramento da Universidade Federal de Viçosa, concluindo-o em dezembro de 2019.

RESUMO

SILVA, Karla Jorge da, D.Sc., Universidade Federal de Viçosa, dezembro de 2019. **Diversidade genética e mapeamento associativo para resistência à fumonisinas em milho tropical utilizando marcadores SNPs**. Orientador: Luiz Antonio dos Santos Dias. Coorientador: Lauro José Moreira Guimarães.

O conhecimento da diversidade genética é fundamental para um programa de melhoramento de plantas pois ajuda entender a relação genética entre as linhagens para direcionar cruzamentos. O mapeamento associativo (GWAS) é um dos principais métodos para relacionar genes e alelos às características de interesse, por meio da co-segregação de marcadores genéticos polimórficos com variação fenotípica. Este trabalho teve como objetivo investigar a diversidade genética de linhagens de milho e verificar a relação entre diversidade genética e padrões heteróticos para a produção de grãos dos híbridos. Com outro conjunto de dados, o objetivo foi identificar regiões genômicas associada com à fumonisina no milho, por meio do mapeamento associativo. Para o estudo da diversidade genética, um total de 1.041 linhagens foram genotipadas por sequenciamento, gerando 32.840 SNPs. Além disso, a diversidade genética de linhagens foi correlacionada com a produção de grãos de 591 híbridos, entre as linhagens genotipadas. No entanto, apenas essas distâncias genéticas entre linhagens não foram suficientes para predizer o desempenho híbrido, uma vez que foi obtida uma correlação de Pearson baixa, embora significativa (0,22, $p < 0,01$) entre distâncias genéticas dos parentais e a produção de grãos dos híbridos. O estudo de mapeamento associativo foi conduzido usando dados fenotípicos de 205 linhagens do painel de milho avaliadas em três ensaios de campo realizados em Sergipe, e um conjunto de 385.654 SNPs. O GWAS foi possível encontrar 45 SNPs significativamente associados à resistência à fumonisinas, agrupados em 19 QTLs. Os QTLs nos bins 2.05, 2.06, 2.08, 2.09, 3.06, 4.05, 5.01 e 10.03 possuem genes candidatos. Com funções preditas implicadas na resistência a patógenos. Os QTLs serão úteis para a seleção assistida por marcadores e para uma melhor compreensão da resistência à fumonisina em milho.

Palavras-chave: Genes. Marcadores moleculares. Melhoramento genético. *Zea mays.*

ABSTRACT

SILVA, Karla Jorge da, D.Sc., Universidade Federal de Viçosa, December, 2019. **Genetic diversity and genome-wide association study for fumonisin resistance in tropical maize with SNPs markers.** Adviser: Luiz Antonio dos Santos Dias. Co-adviser: Lauro José Moreira Guimarães.

The knowledge of genetic diversity is critical for a successful breeding program, helping to dissect the genetic relationship among lines and to identify superior parents. Genome-Wide Association (GWAS) is one of the main methods for relating genes and alleles to the traits of interest, through the co-segregation of polymorphic genetic markers with the genes involved in the variation of the traits. We investigate the genetic diversity of maize inbred lines and verify the relationship between genetic diversity and heterotic patterns based on hybrid grain yield. Subsequently, with another data set, we aimed to genomic regions for fumonisin resistance in maize. For the genetic diversity study, a total of 1,041 maize inbred lines were genotyped-by-sequencing (GBS), generating 32,840 SNPs. Additionally, the genetic diversity of lines was correlated with grain yield of 591 single-cross hybrids. However, solely these genetic distances among inbred lines were not good predictors of the hybrid performance, once a low but significant Pearson's correlation (0.22, $p \leq 0.01$) was obtained between parental genetic distances and hybrids performance. GWAS was conducted using phenotypic data from 205 tropical maize inbreed lines evaluated in three field trials conducted in Sergipe, Brazil and a set of 385,654 high-quality polymorphic SNPs generated using GBS. A total of forty-five SNPs significantly associated with resistance to fumonisin in maize were found and clustered into 19 QTLs. QTLs in bins 2.05, 2.06, 2.08, 2.09, 3.06, 4.05, 5.01 and 10.03 have candidate genes. Genes with annotated functions probably implicated in resistance to pathogens based on previous studies have been highlighted. The QTLs will be useful for marker-assisted selection and for a better understanding of maize resistance fumonisin.

Keywords: Genes. Molecular markers. Plant Breeding. *Zea mays.*

SUMÁRIO

1 INTRODUÇÃO GERAL

O milho (*Zea mays L.* spp.) é uma gramínea da família *Poaceae*, com enorme importância econômica. Possui participação na cadeia alimentar humana e animal, sendo considerado o principal componente na alimentação de aves e suínos, além da fabricação de biocombustíveis, álcool para bebidas e outros fins industriais (Salla et al., 2010). No Brasil, a área total cultivada com milho (primeira e segunda safras) foi de 17,60 milhões de hectares na safra de 2018/19, cuja produção de grãos foi de 98,41 milhões de toneladas (CONAB, 2019). Assim, o Brasil se consolida como um importante produtor mundial deste cereal.

Apesar do grande volume de grãos de milho produzido no país, entre os principais problemas está a baixa qualidade de grãos devido a diversos contaminantes como inseto-pragas de grãos armazenados, fungos, bactérias, micotoxinas e resíduos de agrotóxicos (Lorini e Bacaltchuk, 2007). Entre estes contaminantes, a incidência de grãos ardidos e micotoxinas se destacam, pois afetam a quantidade e qualidade dos grãos, resultando na desvalorização comercial do produto, além de poderem se constituir em barreiras não tarifárias à exportação, podendo causar perdas econômicas, e, mais gravemente, comprometendo a segurança alimentar, visto que as micotoxinas podem provocar diversas doenças em humanos e animais. Os prejuízos causados pelas micotoxinas têm despertado interesse da ciência, do governo e da indústria nacional, visando o aumento do conhecimento e soluções para amenizar os efeitos deletérios causados por essas substâncias (Pereira et al., 2007).

As micotoxinas são metabólitos secundários que comumente contaminam os produtos agrícolas no crescimento da planta ou após a colheita, no transporte e no armazenamento. Esse metábolitos são produzidos por fungos toxigênicos, em condições ambientais favoráveis, inclui muitas espécies do gênero *Fusarium, Aspergillus* e *Penicillium* (Presello et al., 2008). As micotoxinas podem estar presentes, mesmo após a remoção do micélio e uma vez que a maior parte deles são resistentes a tratamentos físicos e químicos, permanecendo nos alimentos durante o processamento e armazenamento (Beber-Rodrigues e Scussel, 2013). As micotoxinas do tipo fumonisinas são produzidas por várias espécies de *Fusarium*, principalmente *Fusarium verticillioides.* Esses fungos estão mais associado ao milho, sendo constatada sua ocorrência natural nessa cultura e produtos derivados em muitas regiões do mundo.

Assim, a definição e caracterização de grupos de diversidade genética gera informações essenciais para o avanço do melhoramento genético. O desenvolvimento de linhagens elite deve ser pautado em combinações genéticas favoráveis dentro de grupos heteróticos definidos e nos cruzamentos entre linhagens de grupos complementares, com o intuito de maximizar a variância

intrapopulacional e as respostas heteróticas nos híbridos finais (Hallauer & Carena, 2009; Munhoz et al., 2009). Assim, a alocação precisa de uma linhagem em um determinado grupo é uma condição importante para a utilização eficiente dos materiais genéticos, direcionando os cruzamentos para formação de híbridos a serem testados e a produção de novas populações fonte para extração de novas linhagens.

Devido ao surgimento de novas raças e a co-evolução entre hospedeiro e patógeno cultivares resistentes precisam ser continuamente desenvolvidas. Essa é uma situação que faz parte da dinâmica da interação planta-patógeno e que precisa ser compreendida pelo melhorista. No processo de transferência de alelos de resistência, os marcadores moleculares do DNA podem ser uma ferramenta bastante útil. Como por exemplo, marcadores ligados aos alelos de resistência, podem ser usados nas etapas iniciais e intermediárias do melhoramento (Yu et al., 2008; Bardol et al., 2013). Nas etapas finais, as inoculações ou exposição da planta ao patógeno em condições de campo são imprescindíveis para confirmar a seleção indireta feita inicialmente por meio dos marcadores.

Muitos caracteres de importância agrícola, como rendimento, qualidade e algumas formas de resistência a doenças, são controlados por muitos genes e são conhecidos como caracteres quantitativos. As regiões dentro dos genomas que contêm genes associados a uma característica quantitativa particular são conhecidas como loco de característica quantitativa (QTLs).

Com os avanços e a redução dos custos das técnicas de genotipagem com marcadores moleculares em alta densidade têm contribuído significativamente para o melhoramento genético. A genotipagem de linhagens fundadoras, progênies derivadas de populações segregantes e painéis de diversidade genética, têm aumentado à eficiência de seleção de genótipos superiores nos programas de melhoramento de plantas. Neste sentido, marcadores SNP (*Single Nucleotide Polymorphism*) têm sido amplamente utilizados em diversos estudos em milho, para várias aplicações, como por exemplo, em análises de caracterização genética entre genótipos e mapeamento associativo apresentando papel importante na priorização de cruzamentos para desenvolvimento de linhagens endogâmicas mais promissoras (Romay et al. 2014; Badu-Apraku et al. 2016; Wu et al. 2016; Zhang et al. 2016; Dari et al. 2018).

Assim, o presente trabalho teve como objetivo realizar um estudo de diversidade genética de linhagens de milho e a relação entre diversidade genética e padrões heteróticos para a produção de grãos dos híbridos. Com outro conjunto de dados, o objetivo foi identificar regiões genômicas e QTLs de alta resolução para resistência à fumonisina no milho.

2 REFERÊNCIAS

Badu-Apraku, B, Fakorede MAB,Gedil M, Annor B, Talabi AO, Akaogu IC, Oyekunle M, Akinwale RO, Fasanmade TY. 2016. Heterotic Patterns of IITA and CIMMYT Early-Maturing Yellow Maize Inbreds under Contrasting Environments. Agron. J. 108:1321–1336.

Bardol, N, Ventelon M, Mangin B, Jasson S, Loywick V, Couton F, et al. 2013. Combined linkage and linkage disequi¬librium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. Theor. Appl. Genet. 126:2717–2736. doi:10.1007/s00122-013-2167-9.

Beber-rodrigues, M. e Scussel, VM. 2013, Mycoflora and Mycotoxicological Quality of Four Freshly Harvested Paddy Rice Cultivars and Relation with Harvest to Industry Reception Timing. Rice Science, 20(4): 303−308, 2013. DOI: 10.1016/S1672-6308(13)60151-1.

Collard, BCY. et al. 2005.An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica, v. 142, n. 1–2, p. 169–196.

CONAB. Companhia Nacional de Abastecimento. Perspectivas para a Agropecuária. v.5 - safra 2017/2018. Brasília: Conab, 2018. Disponível em: <http://www.conab.gov.br/> Acesso em: 28/10/2018.

Dari S, MacRobert J, Ontong M, Labuschagne T. 2018. SNP-based genetic diversity among few-branched- Minnaar-1 (Fbr1) maize lines and its relationship with heterosis, combining ability and grain yield of testcross hybrids. Maydica electronic publication – Vol 63, No 2.

Hallauer AR, MJ Carena MJ. 2009. Maize breeding. In Handbook of plant breeding: cereal breeding.

Lorini, I.; Bacaltchuk, B. 2007. A qualidade desejada na armazenagem de grãos no país. Ambiente em Foco. Disponível em: http://www.ambienteemfoco.com.br/?p=5019. Acesso em: 14 de dezembro de 2019.

Pereira, LA. Impacto de micotoxinas em alimentos no Brasil. Disponível em: http://orbita.starmedia.com/~fitopatologia/micotoxina.htm. Acesso em 10 de dezembro de 2019.

Presello DA, Botta G, Iglesias J, Eyhe´rabide GH. 2008. Effectof disease severity on yield and grain fumonisin concentration of maize hybrids inoculated with Fusarium verticillioides.Crop Prot 27:572–576.

Romay MC, Millard MJ, Glaubitz JC, Peiffer JA, Swarts KL et al. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol 14:R55.

Salla, DA. et al. 2010. Estudo energético da produção de biocombustível a partir do milho. Ciência Rural, Santa Maria, v. 40, n. 9, p. 2017-2022. Disponível em: http://www.scielo.br/pdf/cr/v40n9/a704cr2743.pdf>. Acesso em: 06 jul. 2019

Wu X, Li Y, Shi Y, Song Y, Wang T, Huang Y, Li Y. 2016. Fine genetic characterization o elite maize germplasm using highthroughput SNP genotyping. TheorAppl Genet. 127, 621-631.

Yu, J, Holland JB, McMullen MD, Buckler ES. 2008. Genetic design and statistical power of nested association mapping in maize. Genetics 178:539–551. doi:10.1534/genetics.107.074245.

Zhang, P. et al. 2014. Association mapping for important agronomic traits in core collection of rice (Oryza sativa L.) with SSR markers. PLoS ONE, v. 9, n. 10, p. 16.

Zhang X, Lujiang LL, Hai L, Zhiyong R, Liu D, Wu L, Liu H, JaquethJ , Li B, Pan G, Gao S. 2016. Characterizing the population structure and genetic diversity of maize breeding germplasm in Southwest China using genome-wide SNP markers. BMC Genomics. 17:697.

CAPÍTULO I

HIGH-DENSITY SNP-BASED GENETIC DIVERSITY AND HETEROTIC PATTERNS OF TROPICAL MAIZE BREEDING LINES

VIÇOSA - MINAS GERAIS

2019

APRESENTAÇÃO I

O primeiro capítulo é composto pelo artigo intitulado "*High-density SNP-based genetic diversity and heterotic patterns of tropical maize breeding lines*" que se refere ao artigo apresentado à banca de qualificação. O mesmo foi submetido no dia 12 de julho de 2019 na revista *Crop Science* classificada como A1 no sistema Qualis, e foi aceito no dia 08 de dezembro de 2019.

High-density SNP-based genetic diversity and heterotic patterns of tropical maize breeding lines

Karla Jorge Silva[1], Claudia Teixeira Guimarães[2], José Henrique Soler Guilhen[3], Paulo Evaristo de Oliveira Guimarães[2], Sidney Netto Parentoni[2], Roberto dos Santos Trindade[2], Amanda Avelar de Oliveira[4], Karine da Costa Bernardino[1], Marcos de Oliveira Pinto[2], Kaio Olímpio das Graças Dias[4], Carolina de Oliveira Bernardes[3], Luiz Antônio dos Santos Dias[1], Lauro José Moreira Guimarães[2*], Maria Marta Pastina[2*]

[1]Departamento de Biologia Geral, Universidade Federal de Viçosa, Departamento de Genética e Melhoramento, Viçosa, MG, Brazil. [2]Embrapa Milho e Sorgo, Sete Lagoas, MG, Brazil. [3]Departamento de Genética e Melhoramento, Universidade Federal do Espírito Santo, Alegre, ES, Brazil. [4]Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, Brasil.

*Corresponding authors (lauro.guimaraes@embrapa.br; marta.pastina@embrapa.br)

Abbreviations: Chr, chromosome; CTAB, cetyltrimethylammonium bromide; GBS, genotyping-by-sequencing; Ho, heterozygosity; HG, heterotic group; IBS, identify-by-state; kb, kilobase pairs; Mb, megabase pairs; MAF, minor allele frequency; NJ, neighbor-joining; SNP, single nucleotide polymorphism.

ABSTRACT

Understanding the crop diversity is critical for a successful breeding program, helping to dissect the genetic relationship among lines and to identify superior parents. This study aimed to investigate the genetic diversity of maize inbred lines, and to verify the relationship between genetic diversity and heterotic patterns based on hybrid yield performance. A total of 1,041 maize inbred lines were genotyped-by-sequencing, generating 32,840 quality-filtered SNPs. Diversity analyses were performed using the Neighbor-Joining clustering method, which generated diversity groups. The clustering of lines based on the diversity groups was compared to the pre-defined heterotic groups using the additive genomic relationship matrix and UPGMA. Additionally, the genetic diversity of lines was correlated with yield performance of their corresponding 591 single-cross hybrids. The SNP-based genetic diversity analysis was efficient and reliable to assign lines within pre-defined heterotic groups. However, solely these genetic distances among inbred lines were not good predictors of the hybrid performance, once a low but significant Pearson's correlation (0.22, p-value $\leq$ 0.01) was obtained between parental genetic distances and hybrids performance. Thus, SNP-based genetic distances provided important insights for effective parental selection, avoiding crossed between genetically similar tropical maize lines.

Keywords: *Zea mays* L., genetic variability, breeding program, hierarchical clustering analysis, heterotic groups.

1 INTRODUCTION

Maize (*Zea mays* L.) is a cereal widely used as human food and animal feed, standing out as one of the major grain crops produced worldwide (FAO, 2018). Brazil is the third largest maize producer in the world, with an estimated production of 90 million tons (2018/19) in a total cultivated area estimated in 17 million hectares (CONAB, 2018). To attend the growing demand for food worldwide, maize breeding programs have developed high-yielding cultivars adapted to different environments. However, among several challenges, understanding the genetic extent and structure of the genetic diversity in a breeding program required to meet current production needs under distinct growing conditions, to allow sustained genetic improvement, and to facilitate rapid crop adaptation to adverse effects arising from climate changes and human demographic expansion (Laude and Carena, 2015).

A heterotic group refers to a collection of genotypes (inbred lines or populations) that exhibit similar hybrid performances when crossed with individuals from complementary and genetically distinct germplasm groups (Melchinger et al., 1999; Reif et al., 2004). Inbred lines are often developed from crosses within heterotic groups, whereas promising hybrids are expected from crosses between inbred lines derived from complementary heterotic groups (Hallauer and Carena, 2009; Munhoz et al., 2009). Therefore, it is crucial to understand the genetic diversity within each heterotic group to maximize the exploitation of its genetic variability to develop superior lines.

The advances in molecular marker technologies culminated in several strategies to reveal single nucleotide polymorphisms (SNP), the most abundant polymorphism in eukaryotic genomes with a large potential for high-throughput analysis (Varshney et al., 2009). Genotyping-by-sequencing technology (GBS) (Elshire et al., 2011) improved considerably the generation of SNPs due to its low cost and high-through put generation of polymorphisms along the genomes, being largely used in genetic diversity studies.

Diversity studies based on SNP markers are of great value to assist maize breeders for several purposes (Romay et al., 2013; Badu-Apraku et al., 2016; Zhang et al., 2016; Dari et al., 2018). Molecular characterization of maize inbred lines representing temperate, tropical, and subtropical breeding germplasm from China, Brazil and Mexico showed a substantial level of genetic variation between different breeding pools (Lu et al., 2009). Similar analyses revealed the uniqueness in most of the 450 maize inbred lines widely used in eastern and southern Africa maize breeding programs (Semagn et al., 2012). Recently, 94 early-maturing tropical maize inbred lines genotyped with 15,047 SNP markers were clustered based on their pedigree,

selection history, and endosperm color (Boakyewaa Adu et al., 2019). These and other works confirmed the power of SNP markers for diversity analysis applied in maize breeding programs

In this context, the goal of the present study was to investigate the genetic diversity of maize tropical inbred lines, based on high-density SNP markers, and to verify the relationship between genetic diversity and heterotic patterns based on yield performance of their hybrids.

2 MATERIAL AND METHODS

2.1 Plant Material

The maize diversity panel was comprised of 1,041 inbred lines, belonging to the Embrapa's Maize and Sorghum breeding program, located in Sete Lagoas, state of Minas Gerais, Brazil. These lines were divided, according to their pedigree and breeding information, in three heterotic groups: 530 lines from the Flint group, 357 Dent lines and 154 from an intermediate group, named C, representing lines of several origins, with similar combining abilities when crossed with Dent and Flint testers (Supplemental Table S1).

2.2 Molecular markers data

Genomic DNA was extracted from young leaves based on the cetyltrimethylammonium bromide method (Saghai-Maroof et al., 1984). DNA was quantified and quality checked before sent to the Genomic Diversity Facility, at Cornell University (Ithaca, NY, USA) for genotyping-by-sequencing (GBS, Elshire et al., 2011). GBS was performed using the restriction enzyme *Ape*KI and multiplex formats of 96-plexfor 680 lines (Hiseq2500 1 x 100 bp) and of 384-plex for 380 lines (NextSeq500 1 x 90 bp). The sequences were aligned to the B73 reference genome (AGPv3) using the Burrows-Wheeler alignment (BWA) tool (Li et al., 2009) and the genotype were called using the GBS pipeline available in the software TASSEL-GBS (Glaubitz et al., 2014).

SNPs were filtered for minor allele frequency (MAF) lower than 5%, for maximum of 25% of missing genotypes per locus, and for proportion of heterozygotes per locus inferior to 5%, using the software TASSEL v.5.2.10 (Bradbury et al., 2007). Subsequently, missing genotypes were imputed using the default parameters of the Beagle software (Browning et al., 2016).

2.3 Diversity analysis

MAF and the polymorphic information content (PIC) were estimated for each SNP using the TASSEL v.5.2.10 (Bradbury et al., 2007). PIC tells the discriminatory power of the marker, when considering not only the number of alleles per locus but also its relative frequencies (Botstein et al., 1980), which is expressed by: $PIC = 1 - (\sum_{i=1}^{l} p_i^2 + \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} 2p_i^2 p_j^2)$, where l is the number of alleles per locus; p_i and p_j are the estimated frequencies of the i-thand j-thalleles, respectively.

The genetic distances between maize inbred lines were calculated based on the identity by state (IBS) similarity coefficient (Powell et al., 2010), using the software TASSEL v.5.2.10 (Bradbury et al., 2007). Then, the Neighbor-Joining method was used for clustering analysis (Saitou & Nei, 1987). The number of diversity groups was defined by the K-means, a nonhierarchical procedure (Hartigan & Wong, 1979). The was plot and manipulate clustering, using the package ape (Paradis et al., 2018) and the clustering analysis used the ggtree (Yu et al., 2017). These packages are available in the R software (RCore Team, 2018).

Additionally, the clustering of lines based on the diversity groups was compared to the heterotic groups using the additive genomic relationship (A) matrix of the maize inbred lines calculated according to VanRaden (2008), and clustered with UPGMA (Unweighted Pair Group Method with Arithmatic Mean) using the package ComplexHeatmap (Gu et al., 2016).

2.4 Phenotypic data of hybrids

Field data of maize single-cross hybrids were used to infer about the relationship between genetic distances of parents and their corresponding hybrid performances for grain yield. These hybrids were evaluated between 2006 and 2013, across four locations in Brazil: Campo Mourão-PR, Dourados-MS, Londrina-PR and Vilhena-RO. The experimental trials were arranged in lattice designs with two replicates, using common checks between years and locations. Each experimental plot consisted of two 5 m rows with 0.70 m between rows. Grain yield was determined by weighing all grains in each plot, adjusted to 13% of grain moisture and converted to tons per hectare (ton.ha^{-1}). This breeding historical dataset is highly unbalanced, with an average of 50 genotypes evaluated in each trial.

First, phenotypic analyses were performed for each location-year combination (i.e. environment), considering the effects of replicates and genotypes as fixed, and the effect of blocks within replicates as random. Then, based on the adjusted means of genotypes obtained for each environment, a multi-year and multi-location phenotypic model was fitted, including the effects of years and locations as random, the effect of genotypes as fixed, and the two and

three-way corresponding interaction effects as random. The variance-covariance (**R**) matrix associated to the residual effects within location-year combinations was assumed to be known from the individual analyses per environment as described in Smith et al. (2001).

Out of the 949 evaluated hybrids across environments, 591 had both parents genotyped and were used to fit a GBLUP (genomic best linear unbiased prediction) model, including additive and dominance genetic effects, based on the adjusted means and the weights extracted from the multi-year and multi-location analysis (Munoz et al., 2014; Dias et al., 2019). The additive (**A**) and dominance (**D**) relationship matrices were estimated as described in VanRaden (2008) and Vitezica et al. (2013), respectively.

Pearson's correlation coefficient (r) was used to infer about the relationship between the genetic distances of parents and the adjusted means (g_f),the additive genetic random effect (g_A), the dominance genetic random effect (g_D), and the additive plus dominance genetic random effect (g_{A+D}) of their corresponding hybrids. All linear mixed models were fitted in the ASReml-R v.4 package (Butler et al., 2018).

3 RESULTS

3.1 SNP markers

The GBS sequencing data of 1,041 maize lines generated 904,621 SNP markers, which yielded 32,840 high quality SNP markers after imputation and filtering. The highest and lowest marker densities were observed on chromosomes 5 (17.82 markers per Mb) and 4 (13.55 markers per Mb), respectively. On average, 3,284 SNPs were obtained per chromosome, ranging from 5,189 for chromosome 1 to 2,133 SNPs for chromosome 10 (Table 1).

Table 1. Number, distribution and density of single nucleotide polymorphisms (SNPs) per chromosome of the 1,041 maize inbred lines.

Chr	Size (Mb)	Number of SNP		SNPs/Mb	
		Total	Processed	Total	Processed
1	300.6	135,295	5,189	448.87	17.26
2	237.8	109,951	3,950	462.24	16.61
3	230.8	102,366	3,671	440.85	15.90
4	240.9	98,839	3,263	408.48	13.55
5	217.8	96,770	3,880	444.14	17.82
6	169.2	73,874	2,579	436.18	15.24
7	176.5	76,709	2,908	433.89	16.47
8	175.0	77,889	2,837	444.26	16.21
9	156.4	68,955	2,430	439.34	15.54
10	148.4	63,973	2,133	427.64	14.37
Total		904,621	32,840		

Chr: Chromosome; Processed: filtered for MAF superior than 5%, maximum of 5% of heterozygotes per locus, and imputed for the missing genotypes.

The filtering criteria of MAF greater than 5% was selected to avoid sequencing errors, rare alleles or low coverage loci for a given locus (Oliveira et al., 2018). Out of the total SNP markers, 82% presented MAF between 0.05 and 0.15, whereas 18% presented MAF higher than 0.20. The polymorphic information content (PIC) of the SNP markers ranged from 0.09 to 0.34, with an average of 0.17 (Supplemental Fig. S1).

3.2 Genetic relationship of the maize lines

The clustering analysis based on IBS genetic similarity coefficient and Neighbor-Joining identified twelve distinct diversity groups (G1 to G12) (Fig. 1). The genetic distance between lines ranged from 0.003 to 0.253, with an average of 0.193, and the number of lines within each diversity group varied from 8 to 296 (Fig. 2).

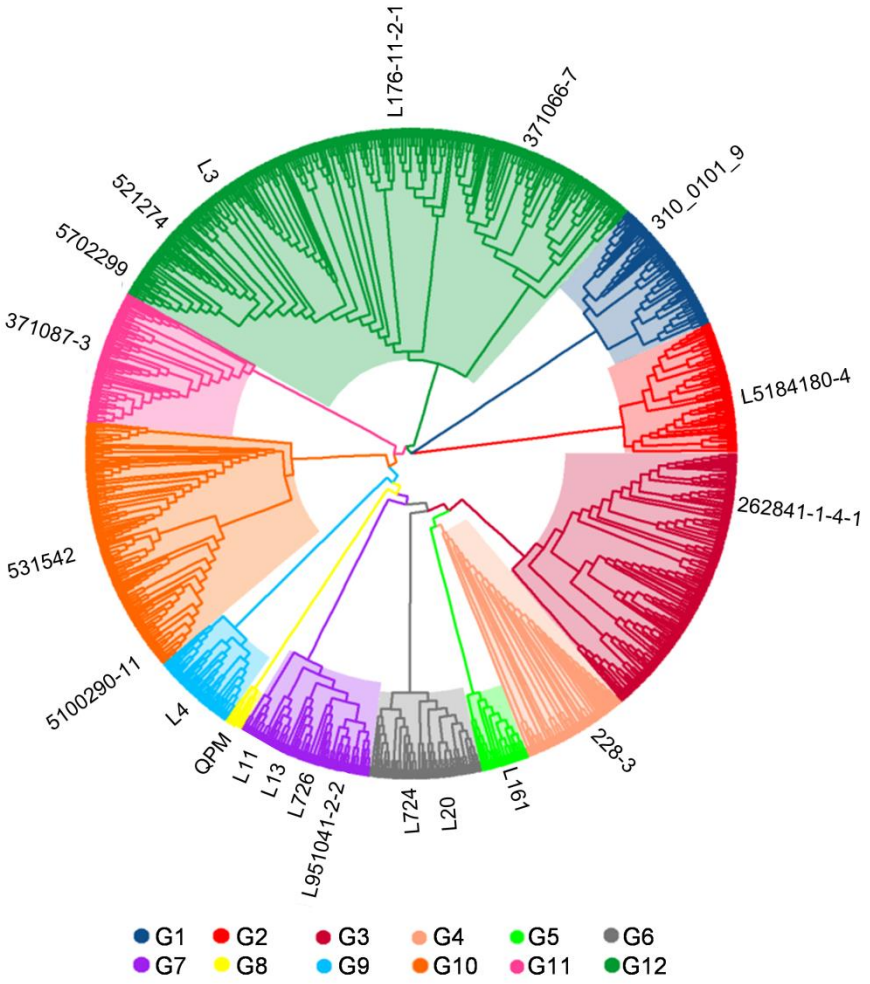


Fig. 1. Dendrogram obtained via Neighbor-Joining using the IBS-derived dissimilarity matrix for the 1,041 maize lines. Different colors indicate the twelve diversity groups.

Most of the twelve diversity groups were composed by more than 60% of lines from a specific heterotic group, Dent, Flint or C (Fig. 2). The groups G1, G2, G10 and G12 represented Flint lines, whereas the groups G3, G4, G5, G9, and G11, included the majority of Dent lines. The G6 group represented C lines, and the groups G7 and G8 included a mixture of lines from all heterotic groups.

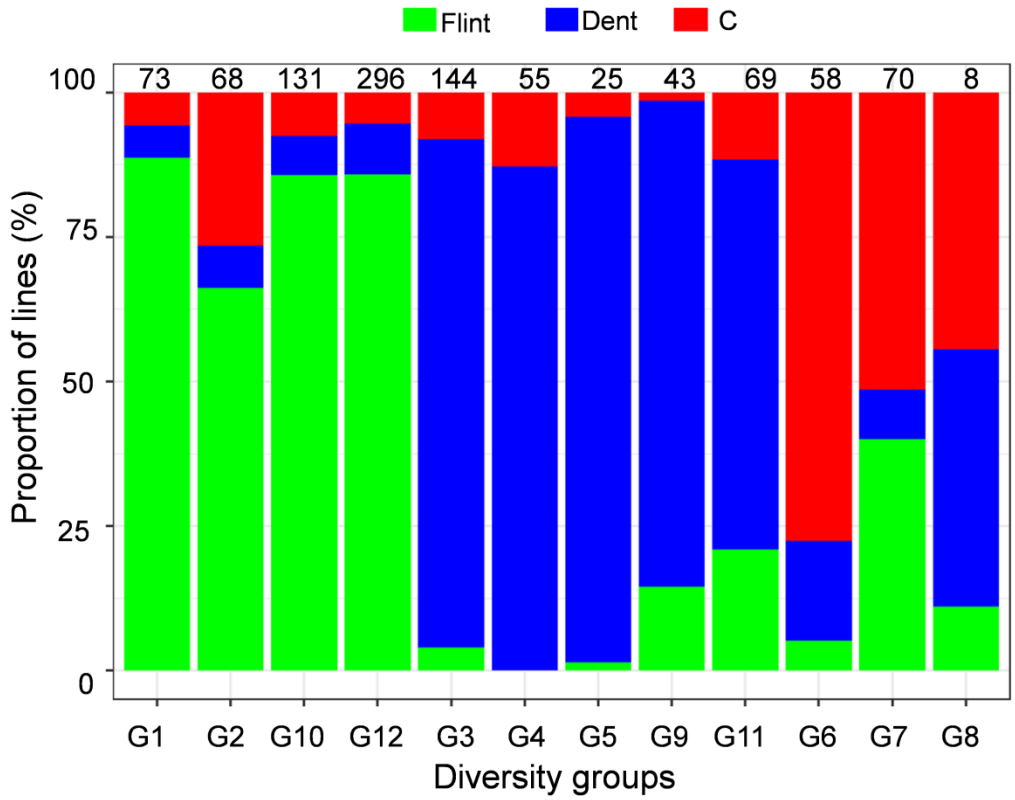


Fig.2. Proportion of Dent, Flint and C lines with in each of the twelve diversity groups (G1 to G12) clustered by Neighbor Joining analysis. Numbers above the bars indicate the number of lines per diversity group.

The A matrix was used to compare the relationship between lines clustered according to the heterotic (Fig. 3a) and to the diversity groups (Fig. 3b). This analysis revealed the existence of subgroups within each heterotic group (Fig. 3a), which were more clear for the Flint and Dent lines than for the C ones. However, the lines within each diversity group were closely related to each other (Fig. 3b), presenting less structure compared to the pre-defined heterotic groups.

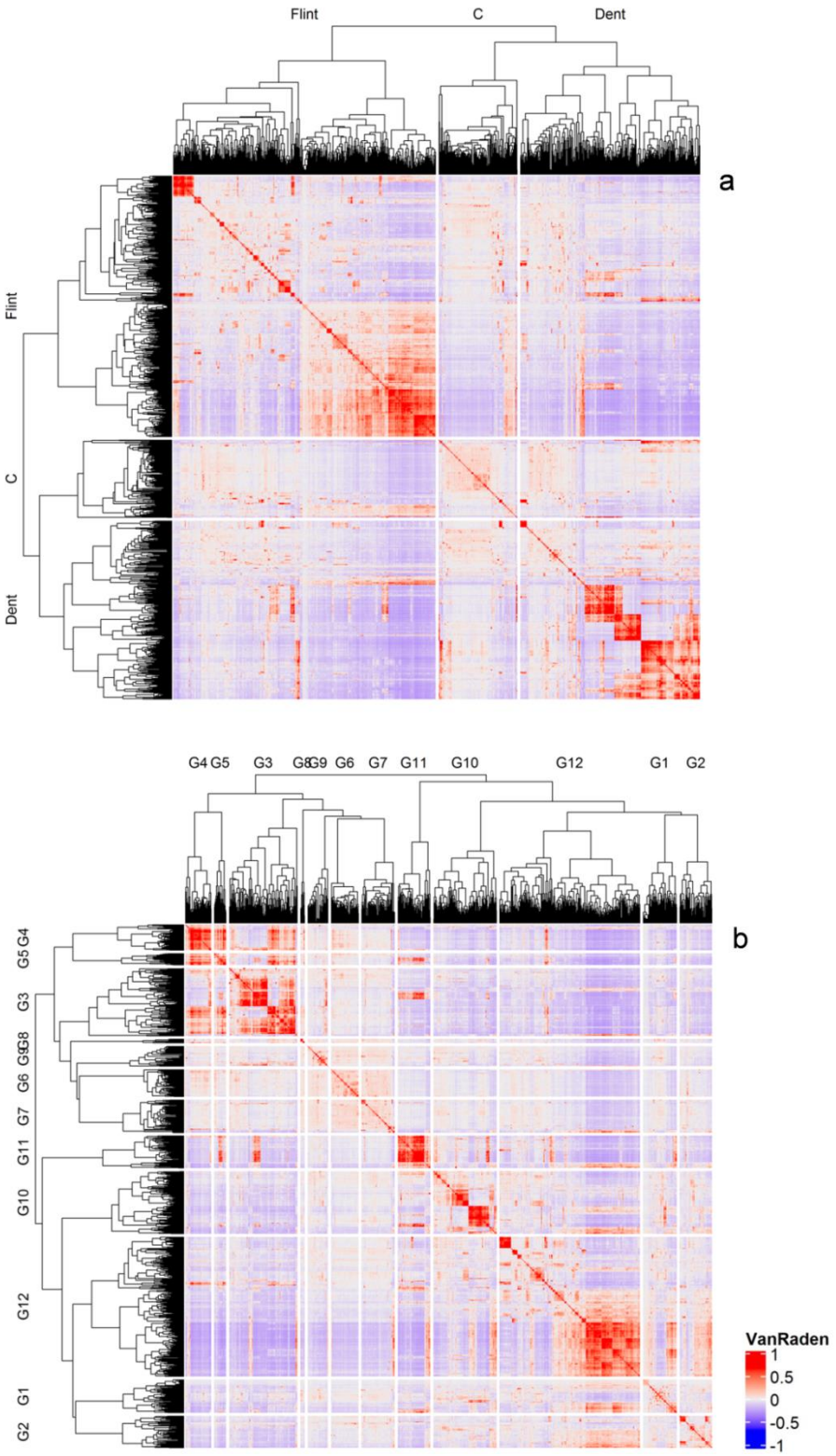


Fig. 3. Heatmap of the additive genomic relationship matrix (A) of 1,041 maize inbred lines. The dendrograms represent UPGMA-clusters, according to: a) predefined heterotic groups (Flint, Dent, and C); and b) the 12 diversity groups identified by the Neighbor-Joining method. Each colored cell represents the additive genomic relationship between lines.

3.3 Correlation between genetic diversity and hybrid performance

There was a significant correlation between the genetic distance of inbred lines ($G._{Dist}$) and the yield performance of their corresponding hybrids, based on the adjusted means of hybrids (g_f, $r^2 = 0.22$) (Fig. 4). As expected, the correlation was higher between the genetic distances of parents and the additive genetic effects (g_A, r = 0.18) than for the dominance genetic effects (g_D, r = 0.13). However, dominance effects slightely increased the correlation when included in the model (g_{A+D}, r = 0.19) (Fig. 4). The scatterplots of the genetic distances of parents ($G.D_{ist}$) and the adjusted means (g_f), the additive genetic effects (g_A), the dominance genetic effects (g_D), and the additive plus dominance genetic effects (g_{A+D}) of their corresponding hybrids are showed below the diagonal in figure 4. The adjusted means of hybrids for grain yield ranged from 3.70 to 8.08 $ton.ha^{-1}$, with an average of 5.90 $ton.ha^{-1}$.

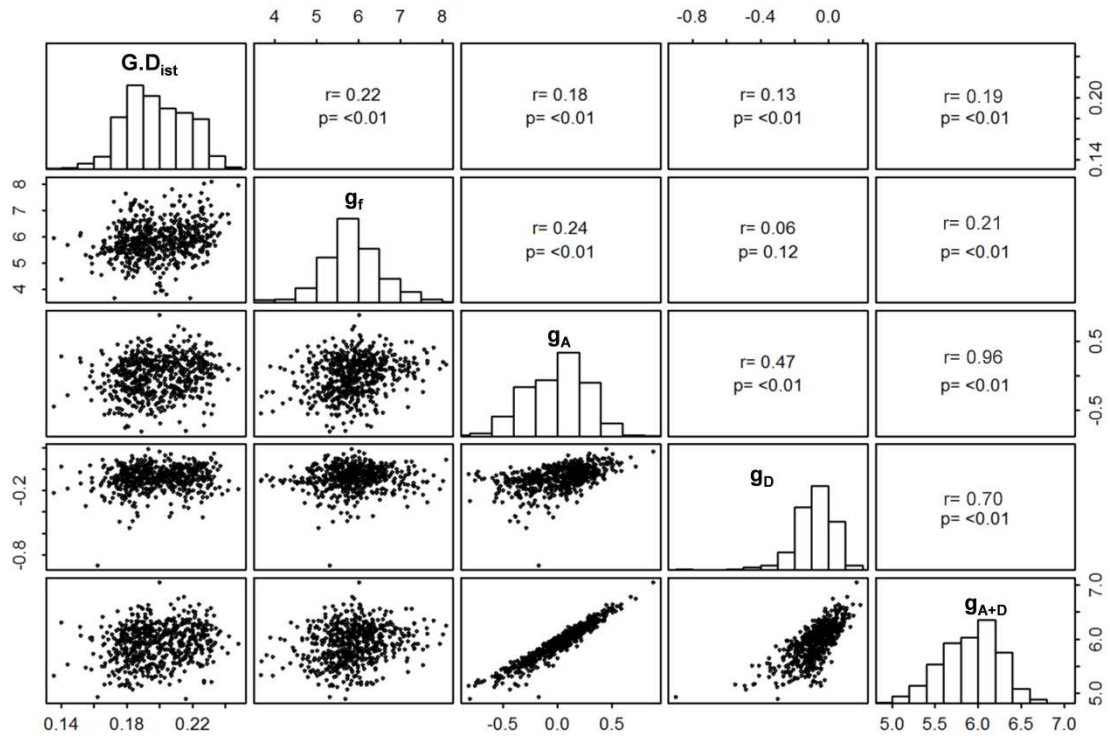


Fig. 4. Scatterplots (lower diagonal) and Pearson's correlations (upper diagonal) between the genetic distances of the parents ($G.D_{ist}$) and the adjusted means (g_f), the additive genetic effects (g_A), the dominance genetic effect (g_D), and the additive plus dominance genetic effect (g_{A+D}) of their corresponding hybrids. Histograms of $G.D_{ist}$, g_f, g_A, g_D and g_{A+D} are presented on the diagonal.

4 DISCUSSION

Genotyping has demonstrated to be a reliable approach to establish the genetic relationships among inbred lines (Govindaraj et al., 2015). In the present study, PIC value ranged from 0.09 to 0.34, with an average of 0.17, which was similar to PIC average of 0.19 in 94 inbred lines genotyped with 15,047 SNPs (Boakyewaa Adu et al., 2019), and PIC values ranging from 0.10 to 0.40 in 128 maize lines genotyped with 143,415 SNPs (Mengesha, 2017). Out of the total SNPs, 82% presented MAF between 0.05 and 0.15 (Supplemental Fig. S1), which agreed with the MAF range of 0.05 to 0.20 of 1,065 SNPs evaluated among 450 lines from CIMMYT (Semagn et al., 2012).

The challenges in analyzing any genetic dataset are to explore whether a given population is homogeneous or contains genetically distinct subgroups (Patterson et al., 2006). The twelve diversity groups generated in our study were compatible with the pedigree information, having one or more elite lines as main founders in each cluster (Fig. 1), confirming the *a priori* information based on pedigree. However, groups composed by large number of lines were structured, showing the existence of genetic variability within diversity groups.

The group G1 consisted of 73 lines, 90% from the Flint heterotic group, highlighting 310-0101-9, an important newly-developed inbreed line. The group G2 was composed by 68 lines, mainly Flint, where the line 5184180-04 stands out for its extra-early maturity as a drought-escape mechanism, which minimizes the exposure to dehydration during the sensitive flowering and post-anthesis grain filling periods. G3 included 144 lines, 97% of them classified as Dent, emphasizing the founder line 262841-1-4-1 that participated effectively as parent of some commercial hybrids.

The G4 grouped 55 Dent lines, represented by the genetic founder L228-3, which probably suggest the narrow genetic basis by its unique branching shape. G5 was composed by 25 lines, mainly from Dent germplasm, having the L161 as an old founder of Dent populations. G6 represented C germplasm with 58 lines, highlighting L20 and L724 that formed two divisions within this diversity group. G7 grouped a mixture of 70 lines from all heterotic groups, which were grouped based on their relationship to the Flint lines L13 and L11, to the C line L726, or to the Dent L951041-2-2. G8 is a small diversity group consisted only by 9 lines, mostly QPM (high protein quality), but with minor contribution in new breeding lines. G9 comprises 43 lines, 88% of which were classified as Dent, including L4, which participated in several hybrids.

The group G10 consisted of 131 lines, including early-maturing line, such as 531542 that was an important Flint founder line, but also newly-developed Flint lines highlighted by 5100290-11 without relationship to other historical Flint lines. G11 consisted of 69 lines, predominantly (84%) from Dent, highlighting the 371087-3 due to its high heterosis when crossed to Flint testers. G12 was the largest diversity group, composed by 296 lines, out of them, 86% were Flint, including the L3, 371066-7 and 521274 that were founders of the Flint heterotic group. Additionally, 26 lines of G12 were classified as Dent, including 5702955 and, only 17 were classified as C, including L176-11-21, which formed subgroups within this diversity group.

The diversity groups were highly consistent with the pedigree information, once lines with common ancestors tended to cluster in the same diversity group, which agreeded with other genetic diversity studies in maize. For example, a clusterization of 770 maize elite lines using 1,034 SNP markers was consistent with their pedigree (Lu et al., 2009). Semagn et al. (2005) divided 450 CIMMYT's maize lines three large diversity groups based on 1,065 SNP markers, which were in general consistent with the pedigree information. Additionally, genotyping 157 important maize elite lines from China with 4,976 GBS-based SNPs offered a clear insight for future exploitation of this germplasm in hybrid development (Leng et al., 2019).

The molecular characterization of maize inbred lines has a strong potential to increase the efficiency of hybrid breeding, once field evaluation is feasible only for a small subset of all possible hybrid combinations. In our study, the correlation between the genetic distance of lines and the grain yield of a subset of their hybrids ($r = 0.22$) had a predominance of additive effects ($r = 0.18$), indicating a high inheritability of this prediction to the next generation. Studying other tropical maize germplasms, Guimarães et al. (2007) described a $r = 0,17$ relationship between genetic distances of inbred lines and hybrid performance. Similar results were also reported by Parentoni et al. (2001), which found a correlation of 0.16 between genetic distances of tropical maize lines and specific combining ability of 378 hybrids from Embrapa based on grain yield across 10 environments. However, the genetic variability of breeding lines within pre-defined heterotic groups can be better explored by high throughput genotyping, which can also be applied to classify new lines into heterotic groups, in order to contribute effectively for high-yielding hybrids development.

5 CONCLUSION

The molecular characterization using high-density SNPs provided important insights about the genetic diversity of maize breeding germplasm, allowing a better understanding of the genetic relationship between breeding lines. Furthermore, the diversity groups were identified within pre-defined heterotic groups and the genetic distance between lines were correlated with the hybrid performance, suggesting an important application of diversity studies in tropical maize breeding programs.

6 REFERENCES

Badu-Apraku, B., M.A.B. Fakorede, M. Gedil, B. Annor, A.O. Talabi, I.C. Akaogu, M. Oyekunle, R.O. Akinwale, T.Y. Fasanmade. 2016. Heterotic Patterns of IITA and CIMMYT Early-Maturing Yellow Maize Inbreds under Contrasting Environments. Agron. J. 108:1321–1336. doi: 10.2134/agronj2015.0425.

Botstein, D.R., R.L. White, M. Skolnick, R.W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-333.

Boakyewaa Adu, G., B. Badu-Apraku, R. Akromah, A.L. Garcia-Oliveira, F.J. Awuku, M. Gedil. 2019. Genetic diversity and population structure of early-maturing tropical maize inbred lines using SNP markers. PLoS ONE 14:e0214810. doi:10.1371/journal.pone.0214810.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S.L. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics Applications Note 23:2633–2635. doi: 10.1093/bioinformatics/btm308.

Browning, B.L., S.R. Browning. 2016. Genotype imputation with millions of reference samples. The American Journal of Human Genetics 98:116–126. doi: 10.1016/j.ajhg.2015.11.020.

Butler, D.G., B.R. Cullis, A.R. Gilmour, B.J. Gogel. 2018. ASReml-R reference manual version 4. Department of Primary Industries and Fisheries, Brisbane.

CONAB. Companhia Nacional de Abastecimento. Perspectivas para a Agropecuária. v.5 - safra 2017/2018. Brasília: Conab. 2018.

Dari, S., J. MacRobert, M. Ontong, T. Labuschagne. 2018. SNP-based genetic diversity among few-branched- Minnaar-1 (Fbr1) maize lines and its relationship with heterosis, combining ability and grain yield of testcross hybrids. Maydica electronic publication. 63:2.

Dias, K.O.G, H.P. Piepho, L.J.M. Guimarães, P.E.O. Guimarães, S.N. Parentoni, M.O Pinto, R.W. Noda, J.V. Magalhães, C.T. Guimarães, A.A.F. Garcia, M.M. Pastina, 2019. Novel strategies for genomic prediction of untested single-cross maize hybrids using unbalanced historical data. Theoretical and Applied Genetics 1432-2242. doi: 10.1007/s00122-019-03475-1.

FAO. 2018. FAOSTAT, Food and agriculture data, http://www.fao.org/faostat/en/#home.

Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire., Q. Sun, et al. 2014. A high capacity genotyping by sequencing analysis pipeline. PLoS ONE. 9:903–916. doi:10.1371/journal.pone.0090346.

Govindaraj, M., M. Vetriventhan, M. Srinivasan. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: An Overview of Its Analytical Perspectives. Genet. Res. Int. 1:1-15. Genetics Research International 2015:431487. doi: 10.1155/2015/431487

Guimarães, S.P., M.E.A.G.Z. Paterniani, R.R. Lüders, A.P. Souza, et al. 2007. Correlation between the heterosis of maize hybrids and genetic divergence among lines. Pesq. Agropec. Bras. 42:811-816. doi:10.1590/S0100-204X2007000600007.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, S.E. Mitchell. 2011. A robust, simple genotyping-bysequencing (GBS) approach for high diversity species. PLoS One. 6:1–10. doi: 10.1371/journal.pone.0019379.

Hallauer, A., M.J. Carena. Maize breeding. In: CARENA, MJ. 2009. Handbook of plant breeding: cereals. New York: Springer.3-98.

Hartigan, J.A., M.A. Wong. 1979. A K-means clustering algorithm. Appl Stats. 28:100–108.

Laude, T.P. and M.J. Carena. 2015. Genetic diversity and heterotic grouping of tropical and temperate maize populations adapted to the northern U.S. Corn Belt Euphytica. 204:661–677. doi: 10.1007/s10681-015-1365-8.

Leng, Y., C. Lv, L. Li, et al. 2019. Heterotic grouping based on genetic variation and population structure of maize inbred lines from current breeding program in Sichuan province, Southwest China using genotyping by sequencing (GBS). Mol Breeding 39: 38. doi:10.1007/s11032-019-0946-y.

Li, H., R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics.25:1754–1760. doi: 10.1093/bioinformatics/btp324.

Lu, Y., J. Yan, C.T. Guimarães, S. Taba, Z. Hao, et al. 2009. Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theoretical and Applied Genetics 120:93-115. doi: 10.1007/s00122-009-1162-7.

Mengesha, W.A., A. Menkir, N. Unakchukwu, S. Meseka, A. Farinola, G. Girma, M. Gedil. 2017. Genetic diversity of tropical maize inbred lines combining resistance to Strigahermonthica with drought tolerance using SNP markers. Plant Breeding, 136:338–343. doi:10.1111/pbr.12479.

Melchinger, A.E. 1999. Genetic diversity and heterosis. In: COORS, J. G.; PANDEY, S. (Ed.). The Genetics and Exploitation of Heterosis in Crops. Madison: American Society of Agronomy; Crop Science Society of America. 99-118. doi:10.1017/S0014479700311085.

Munhoz, R.E.F., A.J. Prioli, A.T. Amaral Junior, C.A. Scapim, G.A. Simon. 2009. Genetic distances between popcorn populations based on molecular markers and correlations with heterosis estimates made by diallel analysis of hybrids. Genetics and Molecular Research 8: 951-962. doi: 10.4238/vol8-3gmr592.

Oliveira, A.A., M.M. Pastina, V.F. Souza, R.A.C. Parrella, R.W. Noda, M.L.F. Simeone R.E. Schaffert, J.V. Magalhães, C.M.B. Damasceno, G.R.A. Margarido. 2018. Genomic prediction applied to high-biomass sorghumfor bioenergy production. Mol Breeding. 38,49. doi:10.1007/s11032-018-0802-5.

Paradis, E. and K. Schliep. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. doi: 10.1093/bioinformatics/bty633.

Parentoni, S.N., J.V. Magalhães, C.A.P. Pacheco, M.X. Santos, T. Abadie, E.E.G. Gama, E. Paiva. 2001. Heterotic groups based on yield-specific combining ability data and phylogenetic relationship determined by RAPD markers for 28 tropical maize open pollinated varieties. Euphytica 121:197-208. doi:10.1023/A:101222112.

Patterson, N, A.L. Price, D. Reich. 2006. Population Structure and Eigenanalysis. PLoS Genet, 2:e190. doi:10.1371/journal.pgen.0020190.

Powell, J.E., P.M. Visscher, M.E. Goddard. 2010. Reconciling the analysis of IBD and IBS in complex trait studies. Nat Publ Gr. 11:800–805. doi: 10.1038/nrg2865.

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen, R.W. Allard. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci. USA 81:8014–8018.

Saitou, N. and Nei M. 1987. The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4: 406–425. PubMed.

Semagn, K., C. Magorokosho, B.S. Vivek, D. Makumbi, Y. Beyene, S. Mugo, B.M. Prasanna, M.L. Warburton. 2012. Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. BMC Genome. 13:113. doi:10.1186/1471-2164-13-113.

Smith, A., B.R.Cullis and R. Thompson. 2001. Analysing variety by environment data using multiplicative mixed models and adjustment for spatial field trend. Biometrics. 57:1138–1147. doi:10.1111/j.0006-341X.2001.01138.

Reif, J.C., X.C. Xia, A.E. Melchinger, M.L. Warburton, D.A. Hoisington, D.L. Beck, M. Bohn, M. Frisch. 2004. Genetic diversity determined within and among CIMMYT maize populations of tropical, subtropical, and temperate germplasm by SSR markers. Crop Science 44:326-334. doi: 10.2135/cropsci2004.0326.

RCore, Team. 2018. R: A language and environment for statistical computing. R Found. Stat. Comput., Vienna.

Romay, M.C., M.J. Millard, J.C. Glaubitz, J.A. Peiffer, K.L. Swarts et al. 2013 Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol 14:R55. doi:10.1186/gb-2013-14-6-r55

Varshney, R.K., S.N. Nayak, G.D. May, S.A. Jackson. 2009.Next generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol. 27:522–530. doi: 10.1016/j.tibtech.2009.05.006.

Vanraden, P.M. 2008. Efficient methods to compute genomic predictions. J. DairySci. 91:4414-4423. doi: 10.3168/jds.2007-0980.

Vitezica, Z.G., Varona, L., Legarra, A. 2013. On the additive and dominant variance and covariance of individuals within the genomic selection scope. Genetics 195:1223–1230. doi: 10.1534/genetics.113.155176.

Yu, G., D.K. Smith, H. Zhu, Y. Guan, T.T-Y Lam. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol. 8:28–36. doi:10.1111/2041-210X.12628.

Zhang, X., L.L. Lujiang, L. Hai, R. Zhiyong, D. Liu, L.Wu, H. Liu, J. Jaqueth, B. Li, G.Pan, S. Gao. 2016. Characterizing the population structure and genetic diversity of maize

breeding germplasm in Southwest China using genome-wide SNP markers. BMC Genomics 17:697. doi:10.1186/s12864-016-3041-3.

7 SUPPLEMENTAL MATERIAL

Supplemental Table S1. Tropical maize inbred lines (1,041) belonging to the Embrapa's Maize and Sorghum breeding program and their corresponding heterotic group, previously defined based on predigree and/or testcrosses, and the genetic diversity groups identified through the NJ clustering analysis.

ID	NAME	Heterotic group	Group
L1	5110314_1	Flint	G10
L2	3130452_8	Dent	G11
L3	L1147	Flint	G6
L4	540942	Dent	G9
L5	L3XL3X1113_01_1_1X52111760_61_01_4	Flint	G12
L6	L5720333_01_3	C	G2
L7	51400358_1	Flint	G10
L8	L521538	Flint	G2
L9	5702063	Flint	G12
L10	5046L3L3_2_9L3_1L3_1_1_1	Flint	G12
L11	5100290_8	Flint	G10
L12	106xL228_3_108_1	Dent	G4
L13	L57500_05	C	G6
L14	L83_3_28	C	G6
L15	3130492_5	Flint	G12
L16	2110863_2	Flint	G12
L17	CMS28_7_1	C	G6
L18	91500196_3	Dent	G5
L19	5110246_14	Dent	G3
L20	3140169_2	Flint	G10
L21	419133_10_1_5	C	G1
L22	412085_20_1	Dent	G12
L23	262841_1_4_1x521236_3	Dent	G3
L24	91500189_1	Dent	G3
L25	91500054_1	Flint	G10
L26	211_0777_5	Flint	G10
L27	211_0646_1	Dent	G5
L28	3120887_2	Flint	G12
L29	51200307_3	Dent	G11
L30	Pasco14xL3xL3_1_1	Flint	G12
L31	3810214_6	Dent	G9
L32	5100259_5	Flint	G10
L33	371060_1	Flint	G1
L34	3130261_6	Flint	G1
L35	3130578_3	Dent	G11
L36	504611_28_2xL3_5	Flint	G7
L37	540927	Dent	G9
L38	3130379_8	Flint	G2
L39	91500163_2	Flint	G12
L40	Pasco14xL3xL3_3_2_2_1	Flint	G12
L41	5702286	Flint	G12
L42	3820975_8	Flint	G2
L43	3130513_4	Flint	G12
L44	L521573	Flint	G1
L45	5100357_5	Flint	G12
L46	L3x280__2_6	Flint	G2
L47	228_3_2xL228_3_1_1	Dent	G4
L48	SP181_71	C	G7

ID	NAME	Heterotic group	Group
L51	106xL228_3_62_2	Dent	G4
L52	552929_D	Dent	G9
L53	3140163_2	Dent	G11
L54	590421_1	Dent	G9
L55	91500171_5	Dent	G1
L56	5702299	Flint	G12
L57	CatetoColombia9671	C	G7
L58	L56_3_228_9560_58	C	G4
L59	C3_58xL3_1_1_1	Flint	G12
L60	410542_3_1	C	G12
L61	310_0150_6	Flint	G11
L62	Pasco14xL3xL3_3_2_2_1_1	Flint	G12
L63	3130464_2	Dent	G3
L64	3150007_8	Flint	G10
L65	3130573_6	Dent	G5
L66	5110259_3	Dent	G3
L67	5110357_4	Flint	G12
L68	3130431_7	Dent	G5
L69	5100248_4	Flint	G10
L70	L3x723726_45XTR6DM25_15	Flint	G2
L71	L176_2_2_1b	C	G12
L72	211_0764_3	Flint	G10
L73	AG75_9_2	Dent	G12
L74	L724	C	G6
L75	3820997_3	Flint	G12
L76	3130292_3	Flint	G1
L77	3811276_4	Flint	G12
L78	L11	Flint	G7
L79	5110395_4	Flint	G12
L80	540749	Dent	G3
L81	3140168_11	Flint	G10
L82	3140153_3	Flint	G12
L83	1017	C	G6
L84	3810227_1	Dent	G5
L85	TR6DM25	Flint	G2
L86	51200063_1	Dent	G3
L87	L56800_48	C	G6
L88	L530939	Flint	G12
L89	3910973_1	Dent	G10
L90	5110368_5	Flint	G12
L91	91500100_1	Flint	G12
L92	3810118_9	Flint	G12
L93	3150001_2	Flint	G10
L94	L540350_11	C	G7
L95	310_0070_11	Flint	G1
L96	L22	C	G5
L97	98_CIM_2_46xL3_1_2_1	Flint	G12
L98	590107_6	Dent	G5
L99	3810216_2	Dent	G9
L100	419403_32_1	C	G12
L101	3130165_6	Flint	G10
L102	L57500_06	C	G4
L103	211_0417_1	Dent	G5
L104	3140070_7	Flint	G12
L105	L57320_1_2	C	G6
L106	L5361_02	C	G7
L107	3140202_5	Flint	G1
L108	420024_19_1	C	G12
L109	L274	Dent	G6

ID	NAME	Heterotic group	Group
L110	L723	Dent	G2
L111	CML344	C	G6
L112	L51502020	C	G10
L113	TR6DM25L311130111_1_3	Flint	G2
L115	288xL3_xL3__1_1	Flint	G12
L116	551614	Flint	G7
L117	3100230_9_1	Dent	G3
L118	guatexL228_3xL228_3_4_1_3_1	Dent	G4
L119	5100257_6	Flint	G10
L120	5110298_1	Flint	G10
L121	3150013_1	Dent	G9
L122	3130529_7	Flint	G1
L123	560032_5	Flint	G10
L124	3130328_8	Flint	G10
L125	C3_56_2	C	G2
L126	3120763_11	Flint	G12
L127	L3XL3X1113_01_1_1X52111760_61_01_11	Flint	G12
L128	530921	Flint	G12
L129	L17_2	C	G6
L130	3100221_7_1	Dent	G3
L131	3130191_7	Dent	G4
L132	371097_4	Dent	G11
L133	Pasco14xL3xL3_3_2_1_1_2	Flint	G12
L134	3150198_1	Flint	G11
L135	51200062_17	Dent	G3
L136	L521254	Flint	G9
L137	L45611_QPM	QPM	G8
L138	L3xL37xL3_3_6_1_1	Flint	G12
L139	91500085_3	Flint	G10
L140	RCNAPCSxL3_586	Flint	G12
L141	590064_1	Dent	G3
L142	5100331_4	Flint	G10
L143	3110727_9	Flint	G2
L144	51200003_5	Dent	G3
L145	3130192_8	Dent	G4
L146	L95035_1_1_23	C	G7
L147	590449_1	Dent	G3
L148	TR6DM25xL3_1_9	Flint	G2
L149	580184_4	Dent	G1
L150	L578034	C	G6
L151	L1835N_1	Flint	G10
L152	PR053RCL540001	Flint	G12
L153	L_L3xL37__5_6_1_1_1	Flint	G12
L154	3140257_15	Dent	G11
L155	L_L3xL37__5_1_1_1_1	Flint	G12
L156	L5133302016_08	Flint	G12
L157	5702914	Dent	G3
L158	3130206_7	Dent	G4
L159	L1835N_6	Flint	G10
L160	C3xD1_133	C	G2
L161	L520992	C	G8
L162	410399_19_1	Dent	G12
L163	5100243_6	Flint	G12
L164	3140189_1	Flint	G1
L165	3810062_3	Flint	G1
L166	531375	Dent	G9
L167	L504611_01xL3_17	Flint	G12

ID	NAME	Heterotic group	Group
L168	202841_1_1_2xL37x33B_9N_3_1_1_B_2_2	Dent	G3
L169	L5780291	C	G6
L170	421236_29_1	C	G12
L171	580073_4	Flint	G10
L172	51400326_11	Flint	G12
L173	L56800_84	C	G6
L174	3150022_4	Flint	G9
L175	5110422_4	Flint	G1
L176	310_0128_6	Flint	G12
L177	51200022_8	Flint	G10
L179	C3_49xL3_1_2_1	Flint	G12
L180	541167	Dent	G3
L181	L3	Flint	G12
L182	211_0795_1	Flint	G12
L183	L96_16_1	C	G6
L184	L8048_7_1_1	Flint	G2
L185	310_0278_11	Dent	G3
L186	CMS61L10-1	Dent	G6
L187	L31113_01_1_1	Flint	G12
L188	3120743_2	Dent	G11
L189	521343	Dent	G5
L190	L_1199x228_3__3_1_1_1_1	Dent	G4
L191	3130616_2	Dent	G5
L192	L530790	Flint	G12
L193	L8048_10_3	Flint	G2
L194	L5046xL3xL3_1_4xL3_1_2	Flint	G12
L195	3150003_1	Flint	G10
L196	5110369_1	Flint	G12
L197	51400365_1	Flint	G10
L198	C3_86_1xL3_2_3_1	Flint	G12
L199	3120714_10	Dent	G4
L200	L18	C	G6
L201	531198	Dent	G3
L202	FP3_9_2	Dent	G8
L203	L85_18_1xL228_3_1_1_1	Dent	G4
L204	RCNAPCSxL228_91	Dent	G4
L205	L22E	C	G7
L206	421549_8_1	C	G10
L207	L5128412891	C	G6
L208	3810077_3	Flint	G2
L209	550707	Flint	G2
L210	L54018_11	C	G7
L211	3130387_2	Flint	G12
L212	RCNAPCSxL3_539A0	Flint	G12
L213	51400290_4	Dent	G11
L214	3820966_1	Flint	G1
L215	51200320_3	Dent	G11
L216	91500104_10	Flint	G10
L217	51200093_12	Flint	G1
L218	211_1074_1	Flint	G12
L219	3130346_1	Flint	G10
L220	L1337_59_6_PH	C	G2
L221	3130402_7	Flint	G1
L222	521162	Flint	G12
L223	5702955	Flint	G12
L224	L504611_28_2xL3_15	Flint	G7
L225	L20	C	G6
L226	211_0606_1	Dent	G5
L227	3130387_6	Flint	G12

ID	NAME	Heterotic group	Group
L228	L5046xL228_3xL228_3_3_5xL228_3_1_2	Dent	G4
L229	91500044_3	Flint	G12
L230	310_0190_3	Dent	G3
L231	L3L31113_01_1_1_1_1	Flint	G12
L232	310_0235_1	Dent	G3
L233	L_228_3x4040__8	Dent	G3
L234	5100155_3	Dent	G3
L235	51200032_8	Dent	G4
L236	5110237_1	Dent	G3
L237	LTR10DM17	C	G7
L238	5100351_9	Flint	G12
L239	3140264_6	Dent	G11
L240	3120923_8	Dent	G11
L241	371066_7	Flint	G12
L243	3140213_3	Flint	G12
L244	723726_45xTR6DM25xL3_17	Flint	G2
L245	421337_36_1	C	G12
L246	419302_2_2	C	G11
L247	5100299_7	Flint	G12
L248	L1027B_31	Flint	G10
L249	L3xL37_3_5_1_1_1	Flint	G12
L250	310_0224_5	Dent	G3
L251	L1199xL3xL3_1_2_1	Flint	G12
L252	3150005_2	Flint	G10
L253	3140112_3	Flint	G10
L254	3150068_2	Flint	G10
L255	L723726_45	C	G2
L256	91500052_8	Flint	G10
L257	5701841	Flint	G12
L258	3120834_3	Dent	G11
L259	Cateto_Al	C	G7
L260	3110727_3	Flint	G2
L261	L1027B_20	Flint	G10
L262	51200052_15	Flint	G12
L263	482071_13	Dent	G11
L264	B2020MF3i	Flint	G7
L265	L513330_01	Flint	G2
L266	560003_3	Flint	G12
L267	482071_18	Dent	G11
L268	L8048_8_2	Flint	G10
L269	L19	C	G5
L270	3150017_10	Flint	G9
L271	51200114_15	Flint	G12
L272	552220_F	Flint	G12
L273	51200079_2	Dent	G11
L274	412196	Dent	G12
L275	262841_1_4_1x514040_2_4	Dent	G3
L276	L228_3FP3RA_1_1	Dent	G4
L277	5110122_1	Dent	G3
L278	51300253_4	Flint	G5
L279	C3_86_1	C	G12
L280	262841_1_4_1	Dent	G3
L281	3110744_3	Flint	G12
L282	5110325_2	Flint	G12
L283	3150002_9	Flint	G10
L284	3120900_7	Flint	G12
L285	310_0246_7	Flint	G1
L286	5110129_4	Dent	G3
L287	L5161_02	C	G7

ID	NAME	Heterotic group	Group
L288	262841_1_4_1L31113_0111_2_5	Dent	G3
L289	5702755	Dent	G3
L290	51400377_5	Flint	G12
L291	L5761_19	C	G6
L292	L521462	Flint	G12
L293	51400374_8	Flint	G12
L294	5064xL228_3_2_1_1	Flint	G7
L295	L5046xL3_5_2_1	Flint	G12
L296	91500058_10	Flint	G12
L297	5100303_13	Flint	G12
L298	LBR190	C	G9
L299	5110312_5	Flint	G12
L300	L951021_2_2	Dent	G7
L301	91500202_1	Dent	G3
L302	310_0090_9	Flint	G12
L303	371045_3	Dent	G12
L304	51200084_1	Flint	G12
L305	3130560_8	Dent	G5
L307	51200003_4	Dent	G3
L308	L53_53_04	C	G7
L309	310_0124_9	Flint	G1
L310	LPF973239	C	G10
L311	CML343	C	G7
L312	51300412_11	Flint	G12
L313	3140126_11	Flint	G12
L314	3130208_3	Dent	G4
L315	50461128_2xL3_8	Flint	G7
L316	310_0268_9	Dent	G3
L317	211_0587_5	Dent	G3
L318	5702973	Dent	G4
L319	5110140_6	Flint	G5
L320	51300308_1	Flint	G12
L321	3130365_1	Dent	G11
L322	51400314_9	Dent	G3
L323	51200384_8	Flint	G1
L324	91500135_3	Dent	G1
L325	5702780	Dent	G3
L326	3130345_5	Flint	G10
L327	L1113_01	Flint	G12
L328	5100318_1	Flint	G1
L329	L19_1	C	G6
L330	3130636_1	Flint	G12
L331	412082_22_1	Dent	G12
L332	5100273_5	Flint	G12
L333	262841_1_4_1x262841_1_8_2_1_10	Dent	G3
L334	L93_DHS_CO	Dent	G3
L335	L3xL37_5_2_1_1_1	Flint	G12
L336	3130441_1	Flint	G12
L337	3140252_9	Dent	G12
L338	211_0812_4	Flint	G12
L339	L5133302014	Flint	G12
L340	3140187_2	Flint	G1
L341	262841_1_4_1L31113_0111_2_3	Dent	G3
L342	3120877_7	Flint	G10
L343	420341_10_1	Flint	G1
L344	L6751F3i	Flint	G12
L345	5702798	Dent	G3
L346	L3x723726_45XTR6DM25_5	Flint	G2
L347	L504611_28_2	Flint	G7

ID	NAME	Heterotic group	Group
L348	3130406_4	Dent	G11
L349	3140125_3	Flint	G12
L350	211_0858_3	Flint	G12
L351	51400315_6	Dent	G3
L352	486xL3xL3_1_1	Flint	G12
L353	310_0231_6	Dent	G3
L354	3150054_2	Flint	G10
L355	L88D_7_2_1	C	G7
L356	451xL228_3xL228_3_1_1	Dent	G4
L357	5110247_1	Dent	G3
L358	5100360_8	Flint	G10
L359	3130513_8	Dent	G11
L360	51200050_18	Flint	G12
L361	590023_1	Flint	G12
L362	371047_2	Flint	G12
L363	L_L3x876518__2_1_1_1_1	Flint	G12
L364	5110290_4	Flint	G10
L365	3811337_3	C	G6
L366	541007	Dent	G9
L367	L540404_11	Flint	G12
L368	310_0246_13	Dent	G3
L369	3810155_5	Dent	G5
L371	5100329_2	Flint	G1
L372	3130459_7	Dent	G3
L373	LPcb14_04B	C	G6
L374	L5780121	C	G6
L375	541011	Dent	G9
L376	51400290_10	Dent	G11
L377	L3xL3X1113_01_1_1X1017XTR6DM25_15	Flint	G2
L378	51300384_11	Flint	G12
L379	310_0253_6	Dent	G3
L380	5110407_8	Flint	G12
L381	580025_1	Flint	G10
L382	412100_4_1	Dent	G12
L383	590101_3	Dent	G5
L384	91500052_9	Flint	G10
L385	91500057_3	Flint	G12
L386	5110376_8	Flint	G12
L387	51200058_8	Dent	G3
L388	5701499	Flint	G2
L389	51400337_8	Dent	G3
L390	51300389_2	Flint	G12
L391	3120843_7	Flint	G1
L392	3140139_10	Flint	G10
L393	371080_4	Flint	G7
L394	3110722_7	Dent	G11
L395	L1199xL3xL3_3_4xL3_1_1	Flint	G12
L396	3140280_8	Dent	G11
L397	590441_6	Flint	G10
L398	310_0248_15	Dent	G3
L399	3110690_15	Flint	G10
L400	91500147_3	Dent	G9
L401	262841_1_8_2	Dent	G3
L402	L5E	C	G4
L403	5110382_5	Flint	G12
L404	L54056_11	C	G6
L405	410288_14_1	Dent	G12
L406	TR6DM25xL3_1_1	Flint	G2

ID	NAME	Heterotic group	Group
L407	L968	C	G6
L408	91500043_6	Flint	G12
L409	CML348	C	G6
L410	590014_3	Flint	G12
L411	91500130_4	Dent	G1
L412	91500053_9	Flint	G10
L413	371101_2	Dent	G11
L414	51200355_8	Flint	G12
L415	51200256_7	Flint	G10
L416	3110731_8	Dent	G3
L417	3140297_12	Dent	G11
L418	310_0068_6	Flint	G12
L419	L56800_27	C	G6
L420	2841	Dent	G3
L421	L31_2_1_2	C	G6
L422	3130229_3	Dent	G3
L423	L_5046_11_28_2xTR10DM17__7	Flint	G7
L424	371049_1	Flint	G12
L425	3110773_8	Flint	G10
L426	51400314_5	Dent	G3
L427	C3xD1105MV	C	G2
L428	560034_1	Flint	G10
L429	482091_10	Dent	G12
L430	311_0781_5	Flint	G1
L431	91500063_7	Flint	G10
L432	550553	Dent	G3
L433	211_0832_3	Flint	G7
L435	5110232_3	Dent	G5
L436	3150079_1	Flint	G11
L437	5110353_3	Flint	G10
L438	L228_3x876518xL228_3_1_2xL228_3_2_2	Dent	G4
L439	3130372_5	Flint	G10
L440	2800652_9	Flint	G12
L441	106xL228_3_158_1	Dent	G4
L442	3810102_1	Flint	G10
L443	L37x3B_7N_3_2_1_B_1_B	Flint	G12
L444	L531003	Flint	G12
L445	L228_3x1612841_102_2__1_4	Dent	G4
L446	51200299_8	Flint	G10
L447	419603_21_1	Flint	G10
L448	5702148	Flint	G12
L449	L1199xL3xL3_3_4xL3_1xL3_2	Flint	G12
L450	3100260_5	Dent	G3
L451	371082_4	Dent	G11
L452	3140316_11	Flint	G12
L453	L6_1_1	C	G6
L454	3150100_5	Dent	G3
L455	CMS61L2	Flint	G7
L456	L6027NxL228_3_7_8	Dent	G3
L457	L202841_1_3_1	Dent	G3
L458	560058_4	Dent	G3
L459	3130402_6	Flint	G1
L460	5110355_2	Flint	G10
L461	5703005	Dent	G3
L462	5110216_2	Dent	G3
L463	L_161x228_3__1_7	Dent	G5
L464	L53	C	G7
L465	L5046	Flint	G7
L466	3140121_1	Flint	G10

ID	NAME	Heterotic group	Group
L467	L521079	Flint	G12
L468	L44	Dent	G6
L469	3130428_1	Dent	G11
L470	590048_11	Flint	G10
L471	GSP50	C	G11
L472	L420	Dent	G6
L473	L37	QPM	G12
L474	L1027B_16	Flint	G10
L475	482081_31	Dent	G9
L476	3140193_2	Flint	G10
L477	3140307_8	Dent	G12
L478	5100291_9	Flint	G12
L479	L5780280	C	G6
L480	51200097_14	Flint	G1
L481	L52111760	Flint	G12
L482	3810113_3	Flint	G11
L483	CMS28_10_2	C	G6
L484	552697_F	Flint	G12
L485	590027_7	Flint	G10
L486	L_228_3x876518__4_1_1_1_1	Dent	G4
L487	L57500_03	C	G7
L488	552651_F	Flint	G10
L489	5702374	Flint	G12
L490	B2020F3i	Dent	G4
L491	3811296_3	Flint	G1
L492	L2_3_2_1	C	G6
L493	5110303_6	Flint	G10
L494	5100280_2	Flint	G12
L495	TR6DM25L311130111_1_10	Flint	G2
L496	L1027B_1	Flint	G10
L497	3120910_2	Flint	G12
L499	L5184180_04	C	G2
L500	590038_1	Flint	G10
L501	371093_3	Dent	G11
L502	51300379_1	Flint	G7
L503	410400_51_1	Dent	G12
L504	3820976_1	Flint	G12
L505	L726	C	G7
L506	L31113_01_1_4	Flint	G12
L507	105xL3_22_3	Flint	G12
L508	5110313_2	Flint	G12
L509	310_0228_10	Dent	G3
L510	228_3_2xL228_3_1_2	Dent	G4
L511	128xL228_3_1_1_1	Dent	G4
L512	482041_37	Flint	G2
L513	3910994_1	Dent	G10
L514	3140228_1	Flint	G12
L515	482041_30	Flint	G2
L516	3120915_6	Dent	G11
L517	5110451_1	Flint	G12
L518	3130411_5	Dent	G4
L519	3820987_1	Flint	G12
L520	262841_1_4_1x262841_1_8_2_1_7	Dent	G3
L521	3150016_7	Flint	G9
L522	L4	QPM	G9
L523	L1027B_23	Flint	G10
L524	310_0064_1	Flint	G1
L525	310_0064_2	Dent	G3
L526	51400315_8	Dent	G3

ID	NAME	Heterotic group	Group
L527	L951041_3_2	Dent	G7
L528	3140290_8	Dent	G11
L529	310_0216_12	Dent	G3
L530	L1154	C	G7
L531	310_0216_11	Flint	G7
L532	3150057_1	Flint	G10
L533	3130407_9	Dent	G11
L534	410453_28_1	Dent	G12
L535	L13	Flint	G7
L536	3810224_1	Dent	G3
L537	5110410_3	Flint	G1
L538	550601	Dent	G3
L539	371083_3	Dent	G11
L540	L56800_43	C	G6
L541	L9003P-1-2	C	G8
L542	F9_03_2_2_1_1	Flint	G1
L543	590444_1	Flint	G10
L544	3810092_4	Flint	G2
L545	L93	C	G3
L546	211_1073_1	Flint	G12
L547	5703027	Flint	G10
L548	5100247_2	Flint	G10
L549	3120664_5	Dent	G11
L550	L5761_80	C	G6
L551	L1199	Dent	G4
L552	51300394_9	Flint	G7
L553	3130181_5	Dent	G11
L554	51200148_14	Flint	G12
L555	3130531_6	Flint	G10
L556	C3_86_1xL3_1_2_1	Flint	G12
L557	211_1048_B	C	G4
L558	Pasco14xL3xL3_1_1_1_1	Flint	G12
L559	5100258_6	Flint	G12
L560	C3xD1_88	C	G2
L561	590435_1	Flint	G12
L563	3130356_9	Flint	G10
L564	3140288_3	Dent	G11
L565	51400332_9	Flint	G12
L566	L1_2_3	C	G6
L567	3120875_10	Flint	G10
L568	3120858_5	Flint	G1
L569	L513330910	Flint	G12
L570	CML340	C	G7
L571	51300421_6	Flint	G7
L572	3100182_10_1	Dent	G3
L573	L_L3xL37__4_6_1_1_1	Flint	G12
L574	L108_I	C	G7
L575	3140070_4	Flint	G12
L576	51200071_11	Dent	G3
L577	L971041_03	Dent	G7
L578	3140257_7	Dent	G11
L579	590085_3	Dent	G3
L580	5110342_2	Flint	G1
L581	560031_6	Flint	G10
L582	580141_3	Dent	G9
L583	310_0160_6	Flint	G1
L584	262841_1_4_1x262841_1_8_2_1_5	Dent	G3
L585	3150005_5	Flint	G10
L586	91500085_2	Flint	G10

ID	NAME	Heterotic group	Group
L587	420628_7_1	Flint	G10
L588	3130165_3	Flint	G12
L589	51400304_1	Flint	G10
L590	419655_24_1	Flint	G7
L591	3110692_2	Dent	G11
L592	C3_55xL3_2_1_2_1	Flint	G12
L593	5110122_3	Dent	G3
L594	429441_36_2	C	G12
L595	91500054_7	Flint	G10
L596	3140277_7	Dent	G12
L597	C100_6	C	G7
L598	L531164	Flint	G9
L599	419377_36_2	C	G12
L600	TR10DM17X1017_1_4	Flint	G7
L601	3150062_4	Flint	G10
L602	5110342_5	Flint	G1
L603	L845	C	G6
L604	3150017_4	Flint	G9
L605	L57500_01	Flint	G12
L606	419083_1_2	C	G10
L607	371085_5	Dent	G11
L608	5100234_4	Flint	G12
L609	580056_8	Flint	G2
L610	3150020_6	Flint	G9
L611	161X26XL228_3_2_7X867_Xa_30	Dent	G5
L612	1199L228_3L228_3_3_1_2_1_1	Dent	G4
L613	310_0182_2	Dent	G3
L614	91500187_2	Dent	G3
L615	2801262_10	Dent	G3
L616	51400313_2	Dent	G3
L617	3150010_6	Dent	G9
L618	410292_6_	Dent	G12
L619	51200071_17	Dent	G3
L620	211_1071_1	Flint	G12
L621	3810215_3	Dent	G9
L622	2800607_9	Dent	G3
L623	211_1050_B	C	G4
L624	371074_9	Flint	G1
L625	5110447_4	Flint	G12
L627	421240_13_1	C	G12
L628	L26	Dent	G3
L629	371108_4	Flint	G1
L630	91500187_1	Dent	G3
L631	3150055_2	Flint	G10
L632	SP054RCL540160_1	C	G2
L633	L1835N_4_2	Flint	G10
L634	L228_3x45611xL228_3_1_1_1_1	Dent	G4
L635	590072_6	Dent	G3
L636	L176_11_2_1	C	G12
L637	429201_1_2	C	G2
L638	L216	Dent	G6
L639	3130226_4	Dent	G3
L640	5100299_12	Flint	G12
L641	521225	Flint	G12
L642	3140318_3	Dent	G11
L643	L9503_27_2_1_2_3	C	G7
L644	51300328_5	Flint	G12
L645	L2891	Dent	G6
L646	3150067_1	Flint	G10

ID	NAME	Heterotic group	Group
L647	310_0182_7	Dent	G3
L648	51400320_5	Flint	G12
L649	3130207_1	Dent	G4
L650	5100262_9	Flint	G10
L651	91500179_3	Dent	G9
L652	410514_30_1	Dent	G12
L653	91500137_2	Dent	G4
L654	3821095_5	Flint	G12
L655	L1199xL3xL3_3_5xL3_2_1	Flint	G12
L656	521236	Dent	G3
L657	521237	Dent	G3
L658	5110437_6	Dent	G11
L659	L1199xL3xL3_3_2_1	Flint	G12
L660	106xL228_3_123_2_1	Dent	G4
L661	211_0495_3	Dent	G3
L662	542161	Dent	G3
L663	L37x3B_8N_1_2_1_1_1_B	Flint	G12
L664	L1837N_8_2	Flint	G12
L665	590413_4	Dent	G3
L666	202841_1_1_2X5761_18_9	Dent	G3
L667	L8_3_1	C	G6
L668	3150045_1	Flint	G10
L669	L1199xL228_3xL228_3_4_3xL228_3_1_2	Dent	G4
L670	L9003P-1	C	G8
L671	5110350_2	Flint	G12
L672	L512388	C	G7
L673	5110150_5	Dent	G3
L674	211_0571_5	Flint	G12
L675	5110411_1	Flint	G1
L676	C3_49xL3_2_2_1	Flint	G12
L677	310_0105_3	Flint	G1
L678	5110326_6	Flint	G12
L679	L52111760_28_01	C	G6
L680	5100330_6	Flint	G2
L681	3140129_3	Dent	G11
L682	310_0071_7	Flint	G1
L683	5100232_1	Flint	G12
L684	371087_3	Dent	G11
L685	590448_7	Flint	G10
L686	429596_42_1	Flint	G10
L687	Pasco14xL3xL3_3_2	Flint	G12
L688	5110185_6	Dent	G3
L689	L5780120	C	G6
L691	2800647_9	Flint	G12
L692	L16	Dent	G6
L693	51300315_9	Flint	G12
L694	3130222_2	Dent	G5
L695	L832187_71	C	G7
L696	228_3	Dent	G4
L697	51200103_14	Flint	G12
L698	410282_4_1	C	G12
L699	521223	Flint	G12
L700	51300405_1	Flint	G2
L701	3120953_9	Dent	G11
L702	3130486_5	Flint	G1
L703	590074_2	Dent	G3
L704	3810085_6	Flint	G12
L705	3120918_3	Flint	G10
L706	3140200_9	Flint	G1

ID	NAME	Heterotic group	Group
L707	310_0236_3	Dent	G3
L708	L421	Dent	G6
L709	560062_2	Dent	G9
L710	TR6DM25L311130111_2_4	Flint	G2
L711	526xL3xL3_2	Flint	G12
L712	202841_1_1_2xL37x33B_9N_3_1_1_B_2_1	Dent	G3
L713	L_373B__3N_1_1_1_B	Flint	G12
L714	57_60_09	C	G9
L715	521274	Flint	G12
L716	L3FP3_4_1	Flint	G2
L717	5701976	Flint	G12
L718	91500048_9	Flint	G12
L719	371067_2	Flint	G12
L720	51200068_1	Dent	G3
L721	161L228_31_7L311130111_1_4	Dent	G5
L722	L57500_08	C	G3
L723	412105_7_1	Dent	G12
L724	L1_2_1	C	G7
L725	5110365_5	Flint	G12
L726	310_0093_9	Flint	G12
L727	262841_1_4_1x514040_1_1	Dent	G3
L728	521283	Flint	G12
L729	410378_25_1	Dent	G4
L730	91500137_5	Dent	G4
L731	3130364_1	Flint	G12
L732	410524_2_2	Dent	G12
L733	L1011	C	G7
L734	L57	C	G6
L735	371056_1	Flint	G1
L736	521280	Flint	G12
L737	543054	Dent	G3
L738	L5720333_02_3	C	G2
L739	L5720333_02_1	C	G2
L740	3130329_9	Flint	G10
L741	5110159_10	Dent	G9
L742	L578073	C	G6
L743	3120740_3	Dent	G11
L744	C3_49xL3_1_1	Flint	G12
L745	98_CIM_2_46	Flint	G12
L746	5100290_11	Flint	G10
L747	91500185_4	Dent	G3
L748	521255	Flint	G12
L749	521256	Dent	G3
L750	L56800_17	C	G6
L751	L65	C	G7
L752	5702805	Flint	G10
L753	L13_1_2	Flint	G6
L755	3130165_7	Dent	G11
L756	L_228_3x4040__19	Dent	G3
L757	211_1075_1	Flint	G12
L758	421421_3_1	C	G1
L759	51300376_6	Flint	G12
L760	LPF02	C	G7
L761	51200287_4	Dent	G11
L762	5702816	Dent	G5
L763	C3_56_2xL3_1_2_1	Flint	G12
L764	C3_49xL3_2_1_1	Flint	G12

ID	NAME	Heterotic group	Group

L765	541095	Flint	G12
L766	541358	Flint	G2
L767	421552_45_1	C	G10
L768	L57330_910	C	G4
L769	5110424_4	Flint	G1
L770	262841_1_4_1x514040_1_7	Dent	G3
L771	51400293_7	Dent	G3
L772	3130487_5	Flint	G10
L773	531542	Flint	G10
L774	L58C1	C	G7
L775	3140153_1	Flint	G12
L776	541361	Flint	G2
L777	371046_1	Flint	G12
L778	L3xL37_1_6_1_1_1	Flint	G12
L779	5110411_2	Flint	G1
L780	3110733_7	Flint	G1
L781	CML347	C	G7
L782	PR053RCL540160_1	C	G2
L783	51400330_12	Flint	G12
L784	531560	Flint	G10
L785	5110367_11	Flint	G12
L786	541109	Flint	G11
L787	B73	C	G7
L788	3120875_3	Flint	G1
L789	91500053_6	Flint	G10
L790	419083_1_1	C	G10
L791	L64	C	G6
L792	541350	Flint	G2
L793	3140120_3	Dent	G11
L794	3140058_2	Flint	G2
L795	3810067_7	Flint	G12
L796	51200067_7	Dent	G3
L797	L6403P_1	Flint	G12
L798	541338	Flint	G10
L799	91500202_4	Dent	G3
L800	541339	Dent	G10
L801	51300395_11	Flint	G7
L802	L1627_54_4	Flint	G12
L803	580092_2	Dent	G4
L804	211_0704_3	C	G12
L805	LPcb04_05	C	G7
L806	3130521_8	Flint	G12
L807	3140289_4	Dent	G11
L808	L4040	Dent	G3
L809	5110382_7	Flint	G12
L810	5110390_3	Flint	G12
L811	3810032_5	Flint	G12
L812	5100274_5	Flint	G12
L813	51200068_4	Dent	G3
L814	310_0101_9	Flint	G1
L815	521237X202841_1_1_2_6	Dent	G3
L816	3110668_5	Flint	G1
L817	91500200_7	Dent	G5
L819	3130428_6	Dent	G11
L820	590418_6	Dent	G9
L821	L563090	C	G8
L822	L56_LBC_3_228_12_02	C	G9
L823	3140247_4	Dent	G12
L824	L5760_022	Dent	G9
ID	**NAME**	**Heterotic group**	**Group**

L825	C3_49	C	G2
L826	311_0771_7	Flint	G10
L827	211_1067_1	Flint	G7
L828	51200112_16	Flint	G10
L829	3810230_1	Dent	G11
L830	410544_39_1	Dent	G1
L831	5701572	Flint	G10
L832	521306	Dent	G4
L833	521549	Flint	G12
L834	371100_2	Flint	G11
L835	3140132_6	Flint	G1
L836	521550	Dent	G2
L837	L5780305	C	G6
L838	3140196_11	Flint	G12
L839	91500069_6	Flint	G12
L840	541163	Flint	G11
L841	L520	C	G7
L842	51200371_8	Flint	G1
L843	211_0824_2	Flint	G12
L844	419440_43_1	C	G12
L845	L3x876518xL3_2_3xL3_1_1	Flint	G12
L846	590024_2	Flint	G12
L847	L2841L228_3_1_9_1	Dent	G3
L848	L951041_2_2	Dent	G7
L849	5100370_2	Dent	G3
L850	541145	Flint	G12
L851	LPF963173	C	G10
L852	2801027_3	Dent	G3
L853	3150033_1	Dent	G12
L854	410403_22_1	Dent	G12
L855	211_1078_1	Flint	G12
L856	3120913_1	Flint	G7
L857	5110468_1	Flint	G10
L858	531360	Dent	G4
L859	161	Dent	G5
L860	482011_43	Flint	G10
L861	5100187_6	Dent	G3
L862	310_0134_4	Flint	G1
L863	310_0134_1	Flint	G1
L864	L5046xL3xL3_3_6_2	Flint	G12
L865	590035_7	Flint	G10
L866	3120755_7	Flint	G12
L867	3150004_13	Flint	G10
L868	3150033_4	Dent	G12
L869	521529	Dent	G9
L870	3120802_6	Flint	G2
L871	521521	Dent	G9
L872	L56800_67	C	G6
L873	590040_10	Flint	G10
L874	521524	Dent	G9
L875	L7051F3i	Flint	G12
L876	5100317_5	Flint	G1
L877	5100220_3	C	G1
L878	3140010_4	Dent	G3
L879	262841_1_4_1x514040_2_5	Dent	G3
L880	310_0243_1	Dent	G3
L881	51300343_8	Flint	G2
L883	91500145_6	Dent	G9
L884	482011_20	Flint	G10
ID	**NAME**	**Heterotic group**	**Group**
L885	51400338_8	Dent	G3

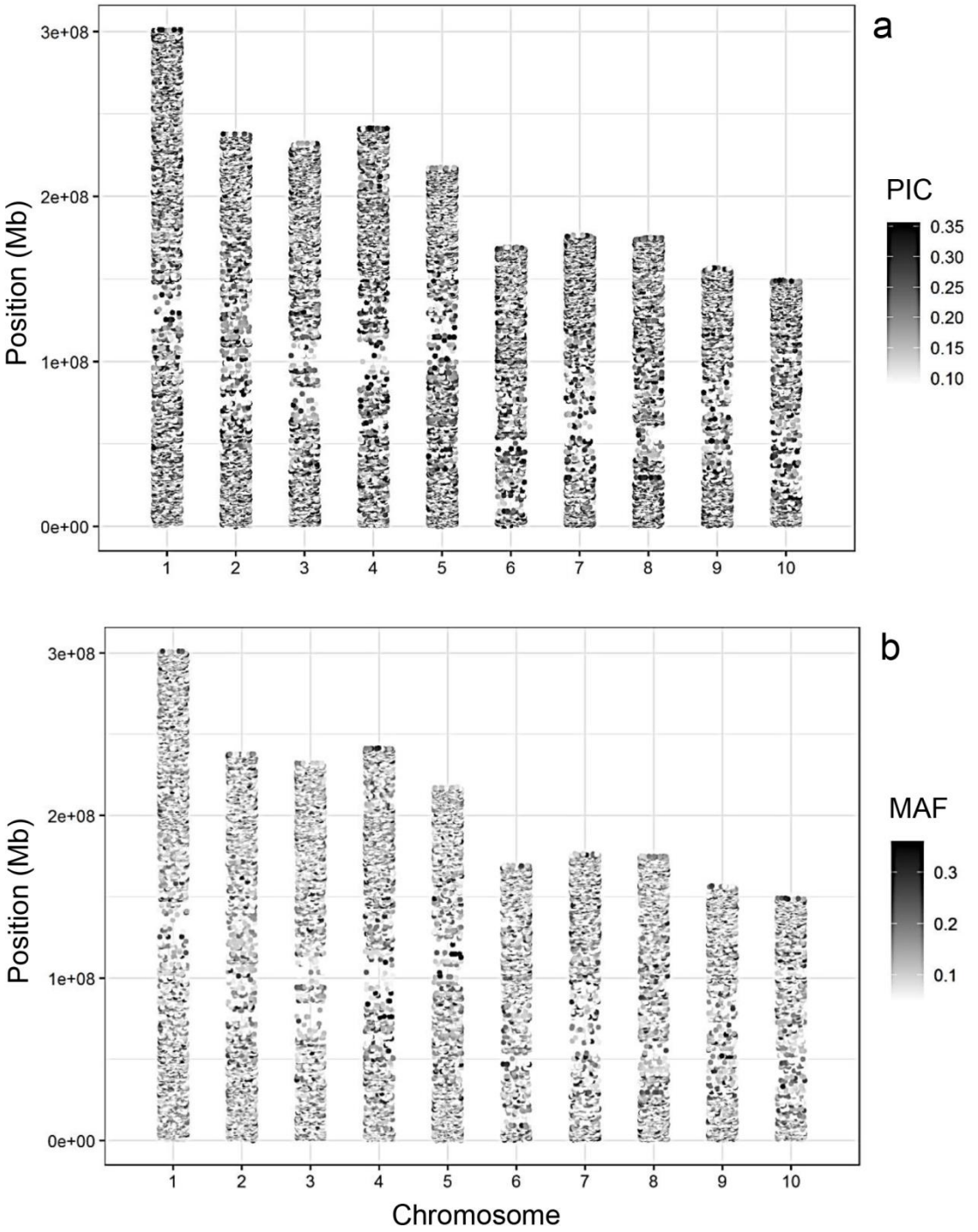
L886	LPF19	C	G2
L887	L_L3xL37_4_2_2_1_1	Flint	G12
L888	3130223_14	Dent	G3
L889	57500_07	Dent	G7
L890	3130460_9	Flint	G1
L891	590011_1	Flint	G12
L892	211_0479_5	Flint	G12
L893	91500125_9	Dent	G3
L894	419427_46_1	C	G12
L895	L228_3_2	Dent	G6
L896	541374	Dent	G10
L897	LPF9_01_2_1_1_1	C	G1
L898	412096_30_1	Dent	G12
L899	410573_53_1	Dent	G10
L900	3110727_4	Flint	G2
L901	448xL228_3xL228_3_1_1	Dent	G4
L902	371102_5	Dent	G11
L903	3140159_5	Flint	G10
L904	412228	Flint	G7
L905	3140310_7	Dent	G11
L906	L228_3x45611xL228_3_2_4xL228_3_1_1	Dent	G4
L907	51300287_10	Flint	G12
L908	5110370_7	Flint	G12
L909	5702223	Flint	G12
L910	410580_45_1	Dent	G1
L911	L5046xL3xL3_2_1_2	Flint	G12
L912	LPF08	C	G7
L913	5110335_2	Flint	G1
L914	530850	Flint	G12
L915	L228_3x45611_5_4_3_1_1	Dent	G8
L916	L3xCMS50_1_18	Flint	G12
L917	L57500_02	C	G3
L918	420561_22_1	C	G12
L919	311_0727_2	Flint	G2
L920	L36	Dent	G2
L921	L228_3XL7003P_10	Dent	G4
L922	419302_2_1	C	G11
L923	5110404_6	Flint	G1
L924	412236	Dent	G12
L925	L5046xL3_1_2_1	Flint	G12
L926	5110404_5	C	G7
L927	L527	Dent	G4
L928	L565218	C	G10
L929	211_1052_B	C	G4
L930	L1170	C	G7
L931	51200365_4	Flint	G10
L932	551303	Dent	G4
L933	3120669_5	Dent	G11
L934	3130387_3	Flint	G12
L935	590100_1	Dent	G3
L936	L56800_79	C	G6
L937	Pasco14xL3xL3_2_1_1_1	Flint	G12
L938	3810145_3	Dent	G4
L939	91500183_8	Dent	G9
L940	3811275_5	Flint	G12
L941	5110404_8	Flint	G1
L942	L5127_33	Flint	G12
L943	FP3_7_1	Dent	G8

ID	**NAME**	**Heterotic group**	**Group**
L944	51200066_3	Dent	G3
L945	51400323_8	Flint	G12

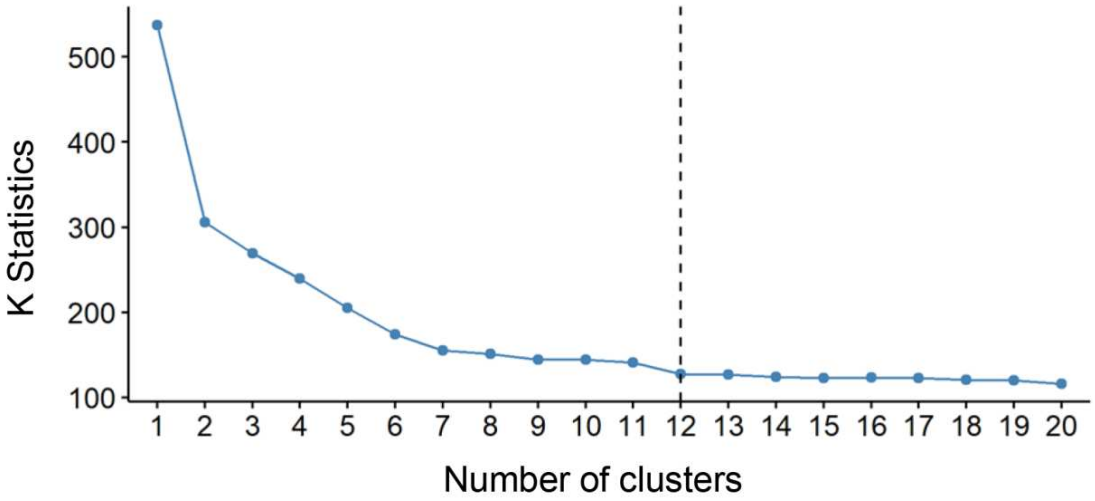
L947	211_0942_2	Flint	G12
L948	580035_2	Flint	G2
L949	51400366_5	Flint	G10
L950	530837	Flint	G12
L951	L57500_04	Flint	G12
L952	L_5046xL3_xL3__3_7_xL3__1_xL3	Flint	G12
L953	550680	Flint	G2
L954	550681	Flint	G12
L955	51200076_6	Flint	G3
L956	51400334_2	Flint	G12
L957	5100304_2	Flint	G12
L958	C3_55xL3__2_1_1_1	Flint	G12
L959	3120714_6	Dent	G4
L960	590013_1	Flint	G1
L961	3150004_8	Flint	G10
L962	51300357_6	Flint	G12
L963	560024_6	Flint	G12
L964	560024_4	Flint	G12
L965	3140310_4	Dent	G11
L966	202841_1_1_2XL228_3X4040_4_9	Dent	G3
L967	310_0165_8	Flint	G1
L968	3150006_5	Flint	G10
L969	5110439_4	Dent	G11
L970	211_0820_3	Flint	G12
L971	L5784180_04	Flint	G2
L972	211_0820_5	Flint	G12
L973	3140279_2	Dent	G11
L974	3140309_7	Flint	G12
L975	5100254_3	Dent	G11
L976	L228_3x45611_5_4_2_1_1	Dent	G8
L977	91500134_10	Flint	G12
L978	3140166_11	Dent	G11
L979	504611_28_2xL228_3_2_1_1	Flint	G7
L980	51300337_9	Flint	G2
L981	L5320708	C	G3
L982	106xL228_3_250_1	Dent	G4
L983	429543_37_1	Flint	G1
L984	3130519_7	Flint	G2
L985	5110247_6	Dent	G3
L986	3811329_1	Flint	G2
L987	5110198_2	Dent	G3
L988	5110153_4	Dent	G3
L989	412204	Dent	G12
L990	429060_3_1	C	G10
L991	3150011_8	Dent	G9
L992	L228_3xC3	Dent	G2
L993	51400340_6	Dent	G3
L994	5110397_8	Flint	G12
L995	3150012_9	Dent	G9
L996	L5046xL3xL3_3_3_2	Flint	G12
L997	429201_1_1	C	G2
L998	91501776	Dent	G9
L999	211_0940_3	Flint	G2
L1000	3140230_13	Flint	G12
L1001	412212	Dent	G12
L1002	590004_1	Flint	G1
L1003	371081_4	Flint	G7

ID	**NAME**	**Heterotic group**	**Group**
L1004	410594_1	Dent	G10
L1005	3130165_4	Dent	G11
L1006	262841_1_4_1x262841_1_8_2_1_8	Dent	G3

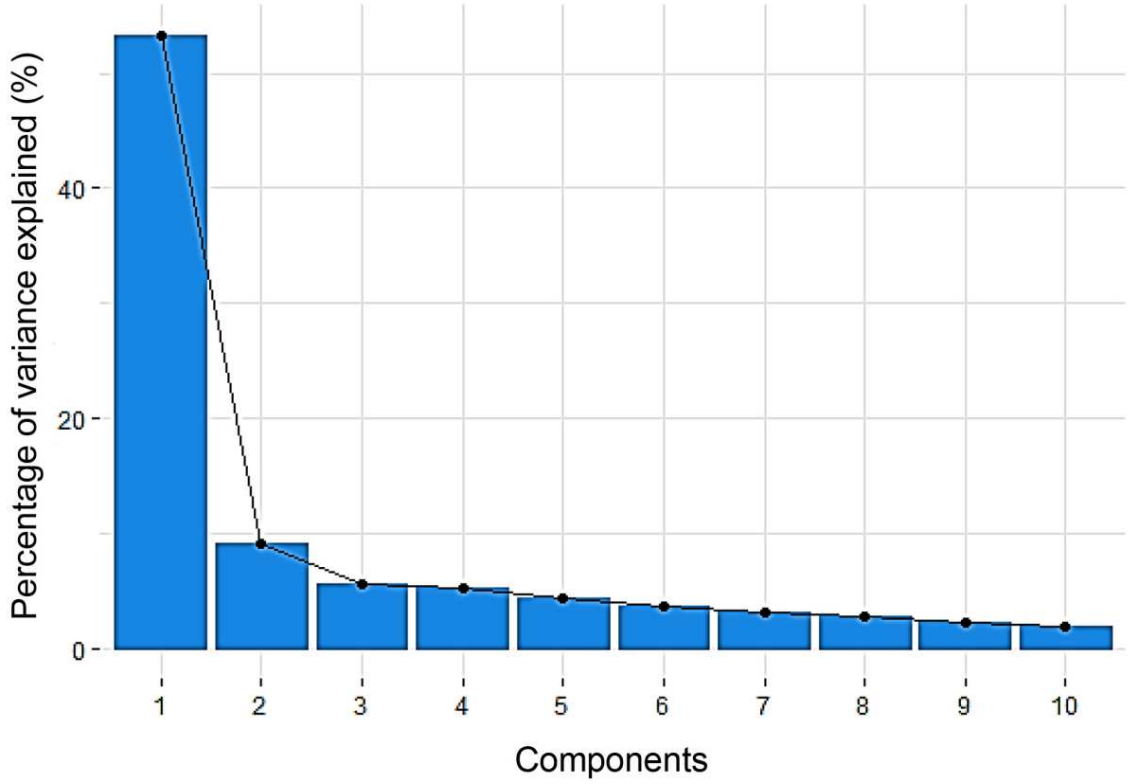
L1007	51200060_3	Dent	G3
L1008	51300415_2	Flint	G7
L1009	L4CT_39_1_1_1_B_1	C	G9
L1011	5702258	Flint	G12
L1012	C3_55	C	G2
L1013	3150025_8	Dent	G9
L1014	482141_19	Flint	G10
L1015	5110305_10	Flint	G10
L1016	L951031_1_1	Dent	G7
L1017	51200052_16	Flint	G12
L1018	3140169_12	Flint	G10
L1019	412220	Dent	G9
L1020	530919	Flint	G2
L1021	262841_1_4_1x262841_1_8_2_1_2	Dent	G3
L1022	CMS61L10-2	Dent	G6
L1023	3820977_3	Flint	G12
L1024	5702751	Dent	G3
L1025	5702778	Dent	G4
L1026	L1835N_9_1	Flint	G10
L1027	L4CT_39_1_1_1_B_2	C	G9
L1028	L3810088_6	Flint	G1
L1029	D1_23_1	C	G2
L1030	3810092_2	Dent	G4
L1031	3811268_1	Flint	G7
L1032	L521163	Flint	G12
L1033	590050_2	Flint	G10
L1034	3810019_2	Flint	G1
L1035	L3x262841_1_4_1_2_1	C	G3
L1036	521275	Flint	G12
L1037	L5046xL3_5_2_1	Flint	G12
L1038	L5702961	Dent	G3
L1039	590040_15	Flint	G10
L1040	504611_01_17	Flint	G12
L1041	3810047_4	Flint	G11



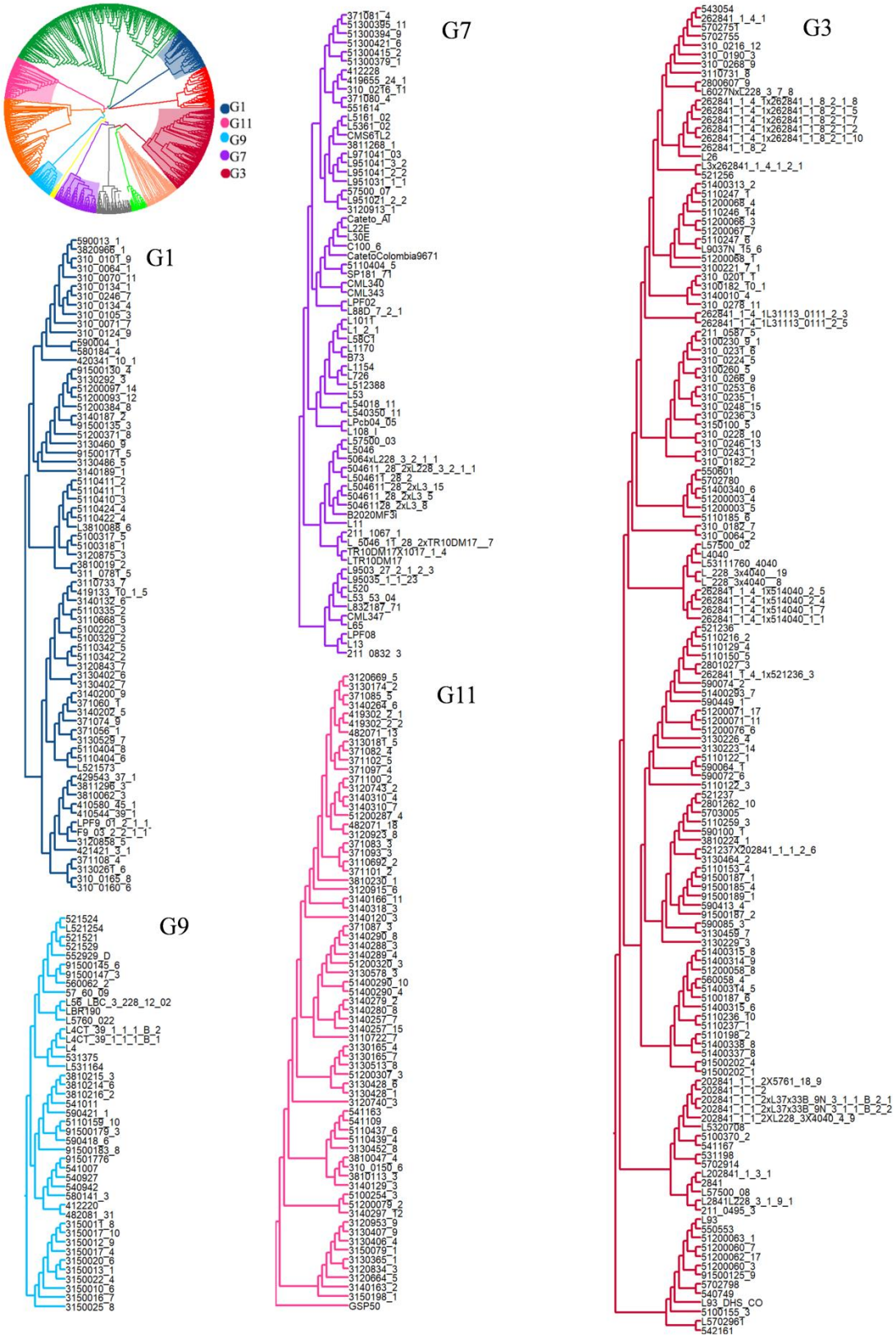
Supplemental Fig. S1. The distribution of PIC, polymorphism information content (a); and MAF, minor allele frequency (b), obtained for 1,041 maize inbred lines. The points on each chromosome represent SNP densities within a window of the 32,840 SNP markers.



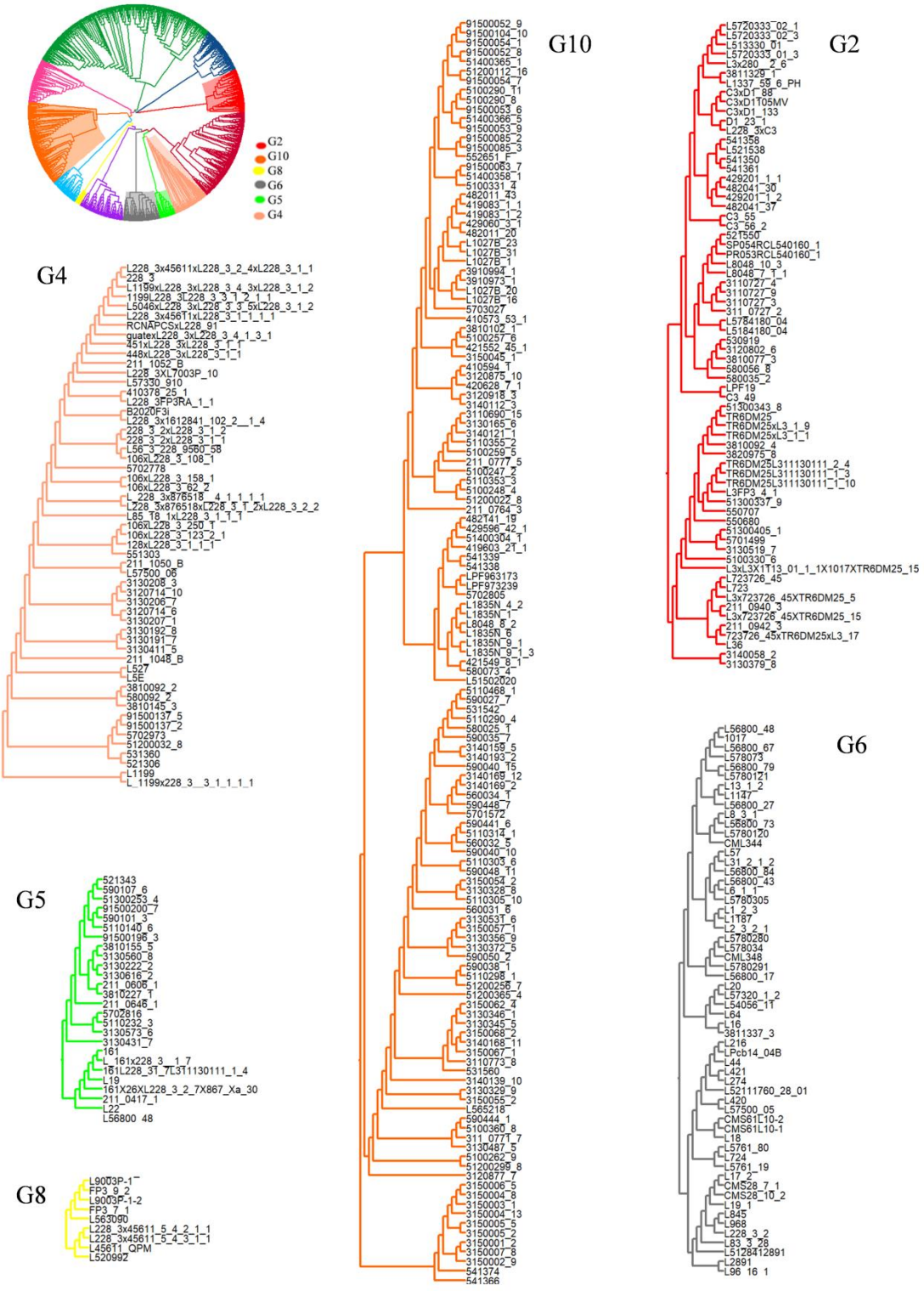
Supplemental Fig. S2 Cut-off point for the definition of the number of diversity groups, using the K-means algorithm, based on the IBS-derived dissimilarity matrix of the 1,041 maize lines



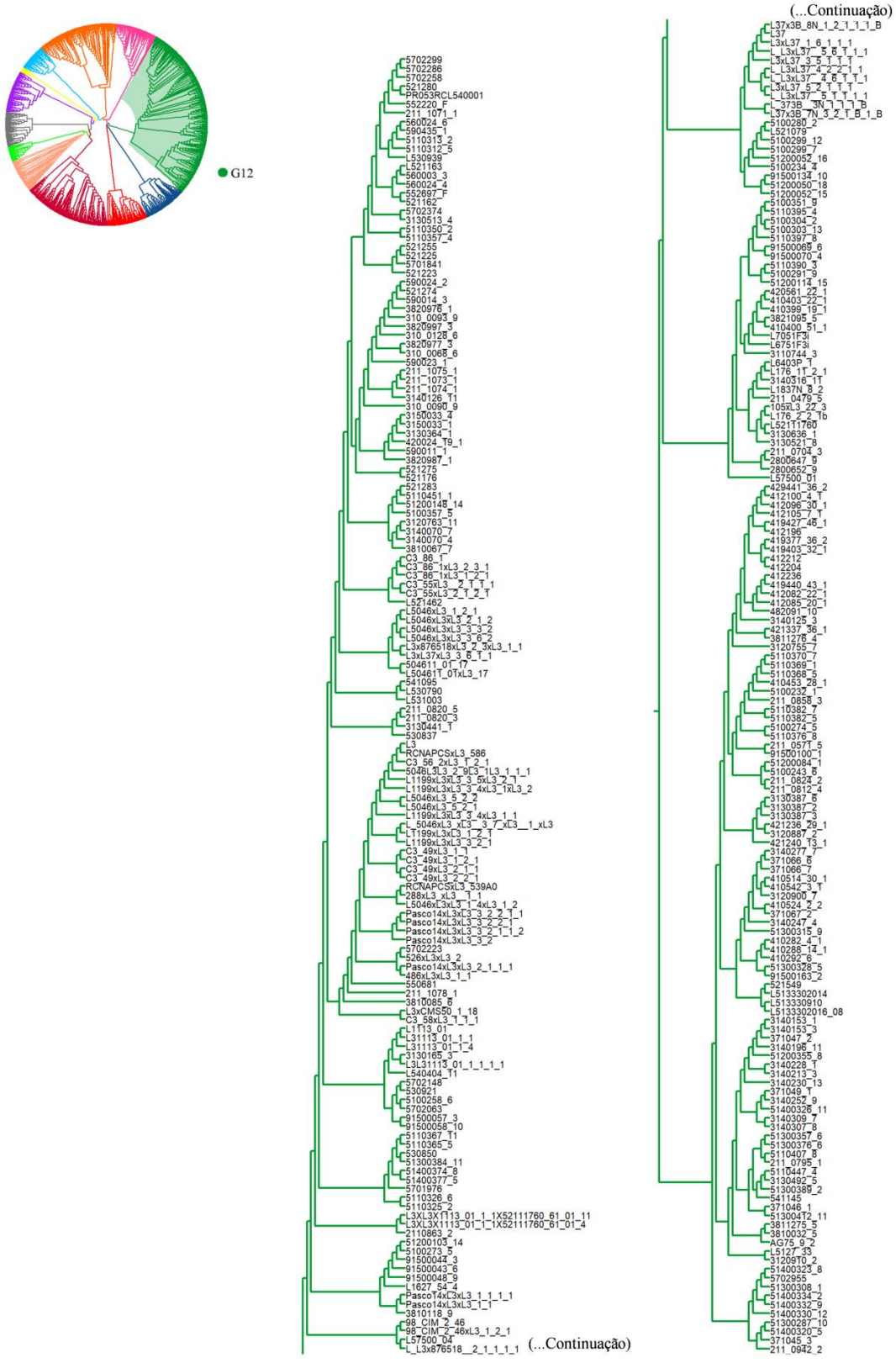
Supplemental Fig. S3 Graphical representation of the variance explained by the eigenvalues of the principal component analysis, using the IBS-derived dissimilarity matrix of the 1,041 maize lines.



Supplemental Fig. S4a Dendrogram obtained by the Neighbor-Joining clustering method based on the IBS-derived dissimilarity matrix of the 1,041 maize lines; the colors indicate the diversity groups G1, G4, G6, G8 and G10.



Supplemental Fig. S4b Dendrogram obtained by the Neighbor-Joining clustering method based on the IBS-derived dissimilarity matrix of the 1,041 maize lines; the colors indicate the diversity groups G2, G7, G9, G11 and G12.



Supplemental Fig. S4c Dendrogram obtained by the Neighbor-Joining clustering method based on the IBS-derived dissimilarity matrix of the 1,041 maize lines; the colors indicate the diversity group G3.

CAPÍTULO II

A GENOME-WIDE ASSOCIATION STUDY FOR FUMONISIN RESISTANCE IN MAIZE

VIÇOSA - MINAS GERAIS

2019

APRESENTAÇÃO II

O segundo capítulo é composto pelo artigo intitulado “A genome-wide association study for fumonisin resistance in maize”. O mesmo após finalizadas as correções será submetido Agronomy Journal.

A Genome-Wide Association Study for Fumonisin resistance in maize

Karla Jorge Silva[1], Claudia Teixeira Guimarães[2], Sylvia Morais de Sousa Tinoco[2], Karine da Costa Bernardino[1], Paulo Evaristo de Oliveira Guimarães[2], Roberto dos Santos Trindade[2], Valéria Aparecida Vieira Queiroz[2], Renata Regina Pereira da Conceição[2], José Henrique Soler Guilhen[3], Natanael Tavares de Oliveira[4], Luiz Antônio dos Santos Dias[1], Lauro José Moreira Guimarães[2], Maria Marta Pastina[2*]

[1]Departamento de Agronomia, Universidade Federal de Viçosa, Viçosa, MG, Brazil. [2]Embrapa Milho e Sorgo, Sete Lagoas, MG, Brazil. [3]Departamento de Genética e Melhoramento, Universidade Federal do Espírito Santo, Alegre, ES, Brazil. [4]Departamento de Bioengenharia UFSJ, São João Del Rei.

*Corresponding author (marta.pastina@embrapa.br)

Abbreviations: Chr, chromosome; CTAB, cetyltrimethylammonium bromide; GBS, genotyping-by-sequencing; GWAS, Genome-Wide Association; Ho, heterozygosity; kb, kilobase pairs; Mbp, mega base pairs; MAF, minor allele frequency; PCA, principal component analysis; QTLs, Quantitative trait loci; SNP, single nucleotide polymorphism.

ABSTRACT

Maize production faces several diseases that cause contamination and are hazardous to health, lowering yield and profit. Thus, plant breeding is one of the most effective and environmentally safe methods to control fungal infection and to increase fumonisin resistance. However, conventional breeding can be hampered by the complex genetic architecture of resistance to fumonisin and marker-assisted selection is proposed as an efficient alternative. The aim of this study was to identify high-resolution QTL (Quantitative trait loci) for resistance to fumonisin in maize. A panel of 205 tropical maize inbreed lines was evaluated in three field trials conducted in Sergipe, Brazil. Fumonisin determination were performed in Embrapa Milho e Sorgo. The lines were genotyped-by-sequencing (GBS), generating 385,654 high-quality polymorphic SNPs. Forty-five SNPs significantly associated with resistance to fumonisin in maize were found and clustered into 19 QTL. The QTLs in bins 2.05, 2.06, 2.08, 2.09, 3.06, 4.05, 5.01 and 10.03 have candidate genes. Genes with annotated functions probably implicated in resistance to pathogens based on previous studies have been highlighted. The QTLs will be useful for marker-assisted selection and for a better understanding of maize resistance fumonisin.

Keywords: *Fusarium verticillioides*. Fumonisin, Resistance. Genome-wide association study. Candidate gene.

1 INTRODUCTION

Maize (*Zea mays* L.) is an important cereal crop grown worldwide for food, feed and processed industrial products, being the third most consumed cereal, after wheat and rice (Chaudhary et al., 2014). Maize production faces several leaf and grain diseases that cause contamination and are hazardous to health, lowering yield and profit. This disease is caused by several fungi that attack and invade developing maize ears and kernels. Fusarium ear rot is mainly caused by *Fusarium verticillioides* (Sacc.) Nirenberg (syn. F. *moniliforme* Sheldon), which reduces grain yield from 10 to 30% (Battilani et al., 2008; Mesterházy et al., 2012) and affects the quality and marketability of grains due to mycotoxin production.

Comprehension of metabolic pathways related to mycotoxins production is essential to target genes related to fungi resistance (Gelderblom et al., 1988). Fumonisin is one of the main toxic metabolites produced by *Fusarium verticillioides.* Fumonisin B1, for example, presents carcinogenic properties (Gelderblom et al., 1996) and is associated with neural tube birth problems in humans (Missmer et al., 2006). Due to hazardous problems to human and animal health it was necessary to define limits for the presence of this mycotoxin in food and feed in most countries (Ferrigo et al., 2016). Tropical countries present more severe infections due to its wet and warm weather, even before harvest. Moreover, wounds caused by insects are readily colonized by fungi that boost infections (Lanubile et al., 2014; Chilaka et al., 2016; Wang et al., 2016). During or shortly after flowering period, contamination with all sorts of Fusarium ear rots is one of the greatest challenges for maize consumption chain. This disease is can occurs, through spores on the silks (Mesterházy et al., 2012).

In solving this problem, some studies were focused on to find markers linked to genes involved in resistance to Fusarium ear rot and/or fumonisin contamination (Chen et al., 2012; Zila et al., 2013; Maschietto et al., 2017; Giomi et al., 2016). For example, a genome-wide association study (GWAS) has been performed for detecting high-resolution QTL for resistance

to fumonisin accumulation for Samayona et al. 2019, used 256 inbred lines in maize and 990,000 SNP markers, resulting in 17 QTL associated with resistance to fumonisin accumulation in maize kernels.

There are currently few studies available on genomic regions and genes associated with fumonisin resistance, partly due to the high correlation between Fusarium ear rot infection and fumonisin contamination as well as the high costs of quantification mycotoxin contamination in large maize populations (Gaikpa et al., 2019). In this present study, GWAS was performed to identify genomic regions associated with fumonisin resistance in tropical maize germplasm. Additionally, tropical maize inbred lines with high levels of fumonisin resistance were selected as important sources for maize breeding programs.

2 MATERIAL AND METHODS

2. 1Genotypes and experimental design

A panel of 205 tropical maize inbreed lines was evaluated in three field trials conducted in Sergipe, Brazil, in a 9 x 9 lattice design, with 81 treatments per trial (75 treatments and 6 common checks), planted into 4 m rows spaced 80 cm apart.

2.2 Fumonisin determination

A sample of 500g of maize grain was finely grounded and a subsample of 10g was used to quantify fumonisin concentration in parts per million, ppm. The analyses were performed in the Laboratory of Food Safety at Embrapa Milho e Sorgo. From this subsample, the fumonisins were extracted in a solution of 100 mL of water/methanol mixture (20/80) and 5g of NaCl in a blender for 1 min. Afterward, it was filtered through Whatman paper and an aliquot of 10 mL of filtered extract was diluted with 40 mL of 0.1% phosphate Tween-20 solution (phosphate buffer). The solution was filtered again with a 1.0 mm microfiber filter, and 10 mL of this

solution was passed through the FumoniTest column. The column was washed with 10 mL of phosphate buffer solution, followed by a second flow of 10 mL of phosphate buffer. The column content was eluted with 1.0 mL of methanol (HPLC grade), collected and mixed with 1 mL of developer. The fumonisin concentration in the grain quantified in samples by the FumonitestTM using the Fluorometer VICAM according to the manufacturer's protocols (VICAM, 2015).

2.3 Estimation of least square means and heritability

Fumonisin data from of the tropical maize inbred lines were first analyzed separately to determine the best fitting model for each trial in Sergipe, and then the best model for the three trials was combined in a single multi-environment trial analysis.

The model was fitted using the statistical package ASReml-R v.3 (Butler et al., 2009), which estimates the variance components using the residual maximum likelihood with the average information algorithm (Gilmour et al., 1995). Diagnostic plots were used to verify the presence of outliers and the fitted models of the residuals. The following mixed model:

$$y_{ijkl} = \mu + e_j + r_{k(l)} + b_{l(kj)} + g_{ijkl} + \varepsilon_{ijkl}$$

where y_{ijkl} is the phenotypic value of the $i^{th.}$ genotype ($i = 1, \dots, I$) from the l block ($l = 1, \dots, l$), in replicate k ($k = 1, \dots, K$), in trial e ($j = 1, \dots, j$); μ is the general mean; e_j is the fixed effect of the trial e; $r_{k(l)}$ is the fixed effect of the replicate k within block l; $b_{l(kj)}$ is the random effect of j^{th} block within replicate k and trial; g_{ijkl} is the random effect of the i^{th} genotype; and ε_{ijkl} is a non-genetic residual effect, with $\varepsilon_{ijkl} \sim N(0, \sigma_\varepsilon^2)$, in which σ_ε^2 is the residual variance.

The heritability (h^2) was estimated using $h^2 = 1 - [\frac{PEV}{2 \times \sigma_g^2}]$, where PEV (prediction error variance) is the mean-variance of the difference between two genetic effects, and σ_g^2 is the genetic variance (Cullis et al., 2006).

2.4 GBS based SNPs

Genomic DNA was extracted from young leaves with the CTBA method (Saghai-Maroof et al., 1984). After extraction, the DNA was quantified on the Qubit® 2.0 Fluorometer, following the manufacturer's instructions (Life Technologies™, EUA), and sent to the Genomic Diversity Facility at Cornell University (Ithaca, NY, USA) for genotyping-by-sequencing (GBS, Elshire et al., 2011). The GBS protocol used the restriction enzyme *Ape*KI and sequenced in a multiplex format of 96-plex for 680 lines (Hiseq2500 1 x 100 bp), whereas 380 lines were performed in 384-plex (NextSeq500 1 x 90 bp). The sequences were aligned to the B73 reference genome (AGPv3) using the Burrows-Wheeler alignment (BWA) tool (Li et al., 2009) and the SNPs were called using the GBS pipeline available in the software TASSEL-GBS (Glaubitz et al., 2014).

Indels and non-biallelic SNPs were removed using the software TASSEL (Bradbury et al., 2007). Missing genotypes were imputed using the default parameters of the Beagle software (Browning et al., 2007), and filtered for MAF lower than 5% and inbreeding lower (F) 0.8 resulting in 385,654 high-quality polymorphic SNPs.

2.5 Population Structure and Estimation of Kinship Matrix

The kinship matrix (K) was calculated by the Identity-by-state approximation, described by Endelman and Jannink (2012). The population structure was determined via principal component analysis (PCA). These analyses were performed using 385,654 polymorphic SNPs with the software TASSEL v.5.2.10 (Bradbury et al., 2007) and *pcaMethods* package in R program (RCore Team, 2018), respectively.

2.6 Genome-Wide Association Study

Four genome-wide association models were tested: (1) “naïve” model using the general linear model (GLM), with no control for population structure and relatedness; (2) GLM

corrected for population structure incorporating the scores for first principal component; (3) GLM corrected for population structure with the scores for first and second principal components; (4) mixed linear model (MLM) with kinship (K) matrix. GLM and MLM models were fitted using the TASSEL version v. 5.2.10 program (Bradbury et al., 2007).

Quantile-quantile plots (Q-Q plots) were used to verify the control for false-positive associations of the tested models. Manhattan-Plot graphs were drawn with physical positions of the markers along the axis X and $-\log_{10}(P)$ on the axis Y. It has been adopted a moderated stringency threshold-based of Samayoa et al. (2019), $-\log_{10}(P)$ equal to 4.0, to consider an SNP significantly associated. R software (RCore Team, 2018) was used to draw the Q-Q plots and Manhattan-Plot graphs.

2. 7 Linkage Disequilibrium (LD) and candidate gene selection

Linkage disequilibrium between locus pairs and LD extension both were determined with the TASSEL v.5.2.10 program (Bradbury et al., 2007). The LD decay was used to investigate candidate genes in disequilibrium with SNPs associated with the target trait, using the MaizeGDB genome browser based on maize B73 genome v3 (RefGen_v3) (Andorf et al., 2010). Annotations of the candidate genes were performed based on BLAST search of the amino acid sequence on the National Center for Biotechnology Information and the BLAST2GO databases (Conesa et al., 2005).

4 RESULTS

The observed mean for fumonisin concentration was 3.70 ppm, ranging from 0.03 ppm (211_0587_5 line) to 8.20 ppm (313_0578_3 line). The heritability (h^2=0.46, considered from low to moderate) was consistent with estimates of previous GWAS studies (Zila et al., 2013; Chen et al., 2016; Giomi et al., 2016).

Table 1 - Variance estimates for fumonisin concentration from 205 maize lines.

Component	Estimate	Std.error	z.ratio
g_{ijkl}	1.79	0.40	4.52
$b_{l(kj)}$	0.31	0.19	1.67
Residual	3.10	0.34	9.06
Heritability	0.46		
Mean (ppm)	3.70		

$b_{l(kj)}$ is the random effect of l^{th} block within replication *k* and trial t; g_{ijkl} is the random effect of the *i-th* genotype. Statistical significance of random effects is calculated using log-likelihood tests.

The degree of relatedness (kinship matrix) among 205 lines was represented by heatmap with six clusterings (colors light blue, orange, pink, red, blue, and green). Some lines that had lower the genetic relationship between them, they without a grouping pattern, therefore were not represented in the six clusterings. Most estimates of kinship remained between 0 and 0.5, on a scale of 0 to 2, being 2 the maximum value of genetic similarity (Figure 1a). Figure 1b presents a biplot of the first two principal components, in which the colored points represent the maize lines clustered according to the six groups identified on kinship heatmap. The first principal component (PC1) explained 9% of the genetic variability, while the second (PC2) explained 3%.

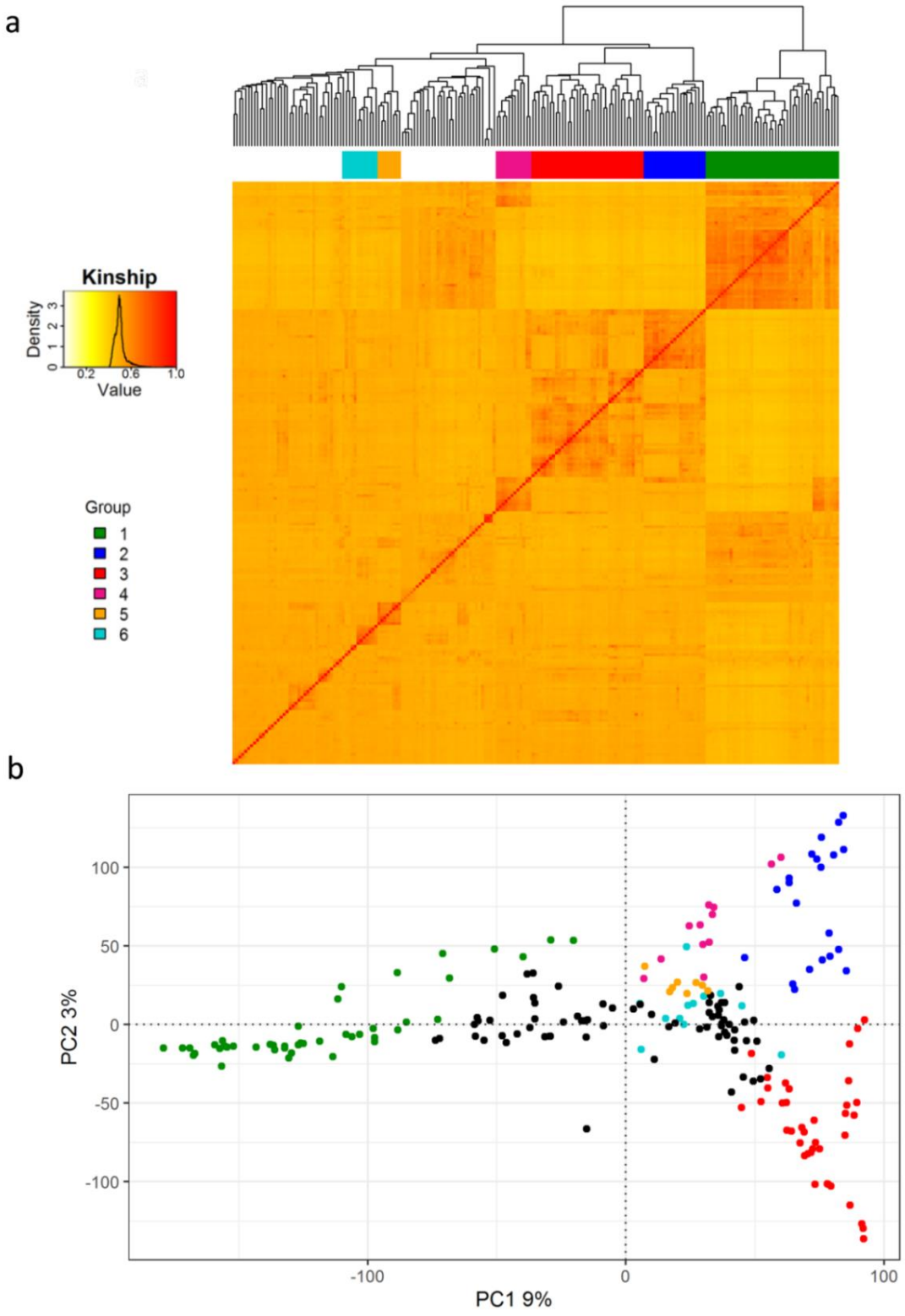


Figure 1 - Heatmap to kinship values for 205 lines and dendrogram based on 385,654 SNPs. (a) The color histogram (kinship) shows the distribution of values for coefficients of coancestry in the kinship matrix, the stronger red bigger the genetic relationship between individuals. *Kinship centered in IBS. (b) Scatter plots of the 205 maize inbred lines according to the first two principal components (PC1 and PC2). Colors represent the genetic diversity groups identified through the histogram (kinship).

The selection of the best model for association mapping was based on the comparison among the distribution of obtained values to $-\log_{10}(P)$ and the expected values for the distribution under the null hypothesis, as represented by the Q-Q Plot (Figure 2). The models corrected for population structure, kinship, or both performed similarly to the naïve model to minimize false positives. Thus, the naïve model was selected for association analyzes for fumonisin resistance. Forty-five SNPs were significantly associated with fumonisin resistance, considering the threshold of $-\log_{10}(P)$ equals 4 (Figure 3). Significant SNPs for resistance to fumosinin were found in bins 1.01, 1.08, 2.05, 2.06, 2.08, 2.09, 3.04, 3.06, 4.05, 4.06, 5.00, 5.01, 10.03 and 10.04 (Table 2).

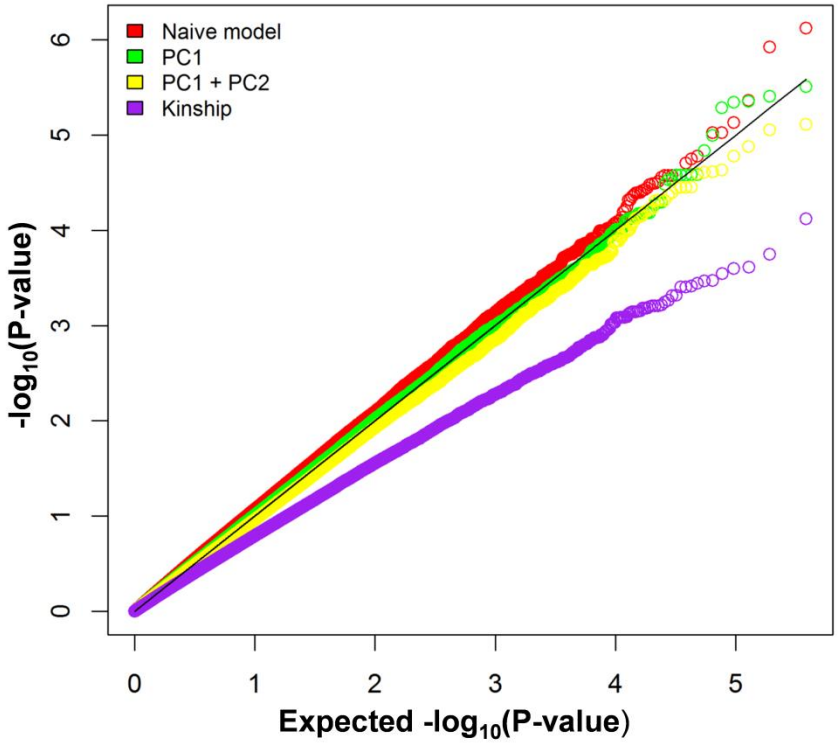


Figure 2 - Quantile-quantile plots (Q-Q plots) for fumonisin association analysis using the "naïve", PC1, PC1 + PC2, and kinship models. The black line is the expected distribution under the nullity hypothesis. Assuming there are few true associations is expected that the observed P values will almost follow the expected P values.

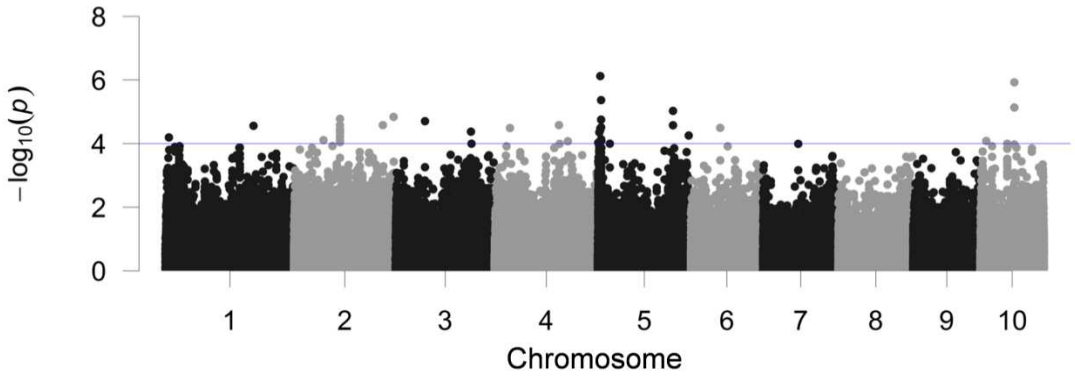


Figure 3 - Manhattan plot of GWAS for contamination by fumonisins in a panel of maize inbred lines. The vertical axis indicates $-\log_{10}$ of P-values, and the horizontal axis indicates chromosomes and physical positions of SNPs.

To investigate candidate genes in disequilibrium with SNPs associated with the fumonisin resistance, we considered the genome-wide LD of this panel, noting that LD tends to decay rapidly to $r^2 = 0.1$ within 10 kb (Supplementary Figure 1). The supporting intervals for the QTL ranged from thousands to millions of bp and were positioned in the B73 genome v3. Significant SNPs were grouped into a unique QTL when located in a genomic region in linkage disequilibrium, resulting in nineteen QTLs for fumonisin (Table 2).We identified 14 genes located within the supporting the interval of each QTL, which were considered as candidate genes for that QTL, as well genes with annotated functions probably implicated in fumonisin resistance based on previous studies (Table 2, Supplementary Table 2).

Table 2 - Single-nucleotide polymorphisms (SNPs) significantly associated with Fumonisin.

QTL[1]	QTL SI[2]	SNP position[3]	Bin[4]	P-value[5]	Candidate gene	Sequence description
1	8533004-8535778	8534120	1.01	0.000064	-	-
2	206561883-206600814	206599838	1.08	0.000028	-	-
3	69040625-69060722	69040625	2.05	0.000078	-	-
4	107622171-107795524	107722261	2.06	0.000061	-	-
		107722317		0.000038	GRMZM2G154156	Ubiquitin ligase complex
		107731269		0.000017	-	-
		107731301		0.000026	-	-
		107781646		0.000048	-	-
		107796371		0.000048	GRMZM2G325815	Domain-containing family protein 2
		107796407		0.000041	-	-
		107910524		0.000040	-	-
		107910552		0.000040	-	-
		107916244		0.000091	GRMZM2G036708	Cysteine synthase
		107916320		0.000036		
		107916792		0.000080		
5	208042001-208045546	208043665	2.08	0.000014	GRMZM2G022213	Zinc finger protein MAGPIE
6	233546104-233549414	233549400	2.09	0.000026	GRMZM2G135013	G-type lectin S-receptor-like serine
7	68754991-69207654	68984340	3.04	0.000065	-	-
8	176845304-176853978	176849366	3.06	0.000042	GRMZM2G060216	Transcription factor LG2
9	36073552-36075444	36074048	4.05	0.000032	GRMZM2G131378	Glycerol-3-phosphate 2-O-acyltransferase 6
10	150664866-150839277	150737434	4.06	0.000026	-	-

(…Continuation)

QTL[1]	QTL SI[2]	SNP position[3]	Bin[4]	P-value[5]	Candidate gene	Sequence description
11	171575219-171576083	171576738		0.000084	-	-
		1040337		0.000094		
12	1039848-3329886	1040343		0.000094	GRMZM2G013200	CRAL-TRIO lipid binding domain
		1040358		0.000094		
13	3327323-3329886	3327619		0.000095	AC210013.4_FGT006	Carbonyl reductase [NADPH] 1
	3403642-3407015	3405488		0.000045	-	-
		4305165		0.000086	GRMZM2G099445	Transcription factor MYB36
14	4305103-5638055	5495018	5.01	0.000038	GRMZM2G104789	Serine/threonine-protein kinase HT1
		5636971		0.000001	-	-
		5637007		0.000033	-	-
		7596002				
		7710528		0.000073	GRMZM2G014618	Sulfate adenylyltransferase cysteine
15	7595319-8419393	7710530				
		7711053		0.000004	GRMZM2G051270	Stress-induced-phosphoprotein 1
		8420393		0.000033	-	-
		176079031			-	-
		176079042			-	-
16	176078031-176082977	176079103	5.05	0.000257	-	-
		176.079105			-	-
17	212795829-212794829	212795829	5.05	0.000057	-	-
19	68779970-68912443	68897845	6.01	0.000032	-	-
18	14449665-14451543	14451414	10.03	0.000083	GRMZM2G083347	NAC transcription factor 29
		80000933		0.000007	-	-
19	79790083-80056274	80000952	10.04	0.000001	-	-

[1]QTL: QTL coding; [2]QTL SI for the supporting interval of the QTL on the RefGenB73_v3 region in which appreciable linkage disequilibrium is observed between SNPs.[3]SNP position for the position in bp of the significant SNP on the RefGenB73_v3; [4]Bin in which QTLs are located; [5]P-value for the association between polymorphic variation at the SNP and phenotypic variation for fumonisin.

The frequency of alleles related to fumonisin resistance was estimated for nineteen significantly associated SNPs, for 10% of lines more resistant (top) and the 10% of lines more susceptible (bottom) (Figure 4A). Figure 4B shows the highest frequency of favorable alleles (RR - alleles that increase the resistance) in top lines and, the highest frequency of unfavorable alleles (SS) in bottom lines. The mean of favorable alleles for the more resistant inbred lines (top 10%) was 0.573, while the mean favorable alleles for less resistant inbred lines (bottom 10%) was 6,88 (Supplementary Table 1).

Some lines from the top group were derived from founding dent lines (L-228-3x45611x228-3-2-4, L-161x228-3-1-7, and L2280-31612841-102-2-1-4), as well as parents of older Embrapa's hybrids (L20 and L724). Inbred lines showing minor alleles had the highest fumonisin frequency (Supplementary table 1, Figure 4C), being the mean 10% of lines more resistant of 0.54 ppm and 10% of lines more susceptible of 6.88 ppm.

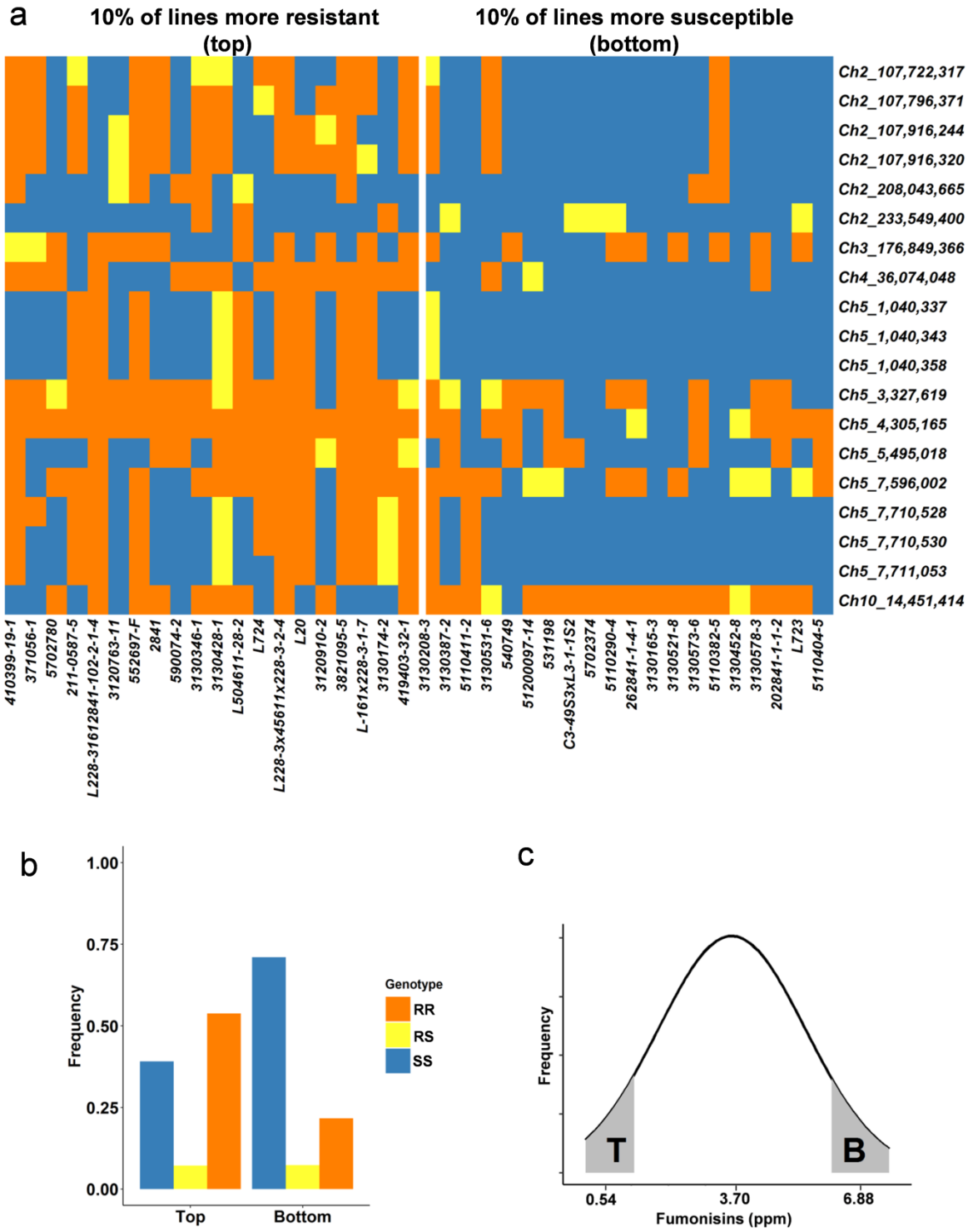


Figure 4 - a) Heatmap with single-nucleotide polymorphisms (SNPs) significantly associated with fumonisin and the respective maize lines that are resistant (top) and susceptible (bottom). b) Frequency of favorable alleles and corresponding genotypes. c) Dispersion of fumonisin data.

5 DISCUSSION

The inbreed maize lines showed significant differences and genetic heritability from low to moderate for fumonisin resistance, which indicate that there is additive genetic variability for this trait. The heritability was similar to those reported by Samayoa et al. (2019), for 270 inbred lines maize to the fumonisin content (h^2 = 0.42). Low heritability for trait emphasizes the importance of implementing marker-assisted selection methods based on stable QTLs in order to increase resistance fumonisin. In this context, genome-wide association would be even more efficient than selection programs based on the phenotype.

The correlations between Fusarium ear rots resistance and fumonisin ranging from 0.87 to 0.92 in some studies (Bolduan et al., 2009; Robertson et al., 2006), selection for resistance to Fusarium ear rot has been proposed as a simpler method to reduce indirectly contamination with fumonisins (Löffler et al., 2010). However, Eller et al. 2008 realized selection for resistance to Fusarium ear rot and defined that selection for reduced Fusarium ear rot could have limited effectiveness to improve resistance to fumonisin. Accordingly, more QTL studies to detect specific genomic regions involved in resistance fumonisins are needed.

Romay et al. (2013) (r^2=0,1;1 kb), Zila et al. (2013) (r^2=0,2; 10 kb), and Zila et al. (2014) (r^2=0,2; 1 kb) reported distinct values for linkage disequilibrium decay in maize. Lu et al. (2010), through haplotypes reconstruction with 2,052 SNPs markers for 305 inbred lines, identified an LD average extension of 10 kb. Mean LD r^2 between the chromosome and SNPs dropped below 0.1 within approximately 10 kb (Supplementary Figure 1). This value for LD is observed because this is a measure dependent on a particular group of materials and a region of the genome with differences for the median values of r^2 (Romay et al., 2013). According to Lu et al. (2010), haplotype-based analysis

substantially improved the efficiency of LD mapping, thus provide a good basis for association mapping regarding fumonisin resistance.

Forty-five SNPs significantly associated with fumonisin resistance were grouped in 19 high-resolution QTLs. Genes at 10 kb of distance or on inferior distances to positions these QTLs were considered candidate genes and were characterized according to the MaizeGDB genome browser. Only genes with annotated functions to resistance to pathogens were discussed. The QTLs in bins 2.05, 2.06, 2.08, 2.09, 3.06, 4.05, 5.01 and 10.03 have candidate genes. Genomic regions 2.05, 2.08, 3.06, 4.05 and, 10.03 were associated with fumonisin in other studies (Zila et al., 2014; Chen et al., 2016; Maschietto et al., 2017; Coan et al., 2018; Samayoa et al., 2019) (Supplementary Table 3).

There are different immune strategies for defending against pathogens in plants (Wit et al., 2007). Pathogen-associated molecular patterns (PAMPs), for example, are patterns recognized by receptors (pattern-recognition receptors - PRRs), which induce the immune response (pattern triggered immunity - PTI). PRRs can be categorized as receptor kinases localized on plasma membrane (RKs) or receptor-like proteins (RLPs) (Boutrot and Zipfel, 2017; Zhang et al., 2017), which reinforces the host defenses. The effector-triggered immunity (ETI), mediated by resistance proteins (RPs), is a secondary immune response that when activated allows the plant to stop the pathogen development. During the induction of local immune responses, a systemic acquired resistance (SAR) can become activated. The maize, for example, when infected by *Fusarium verticillioides*, expresses a set of defense genes (Lanubile et al., 2014; Wang et al., 2016). This response seems to play a primary role in the resistance of maize to *Fusarium verticillioides*, where salicylic acid and jasmonic acid signaling pathways can be involved (Wang et al., 2016). Hence, genes directly involving in the immune response in plants are more suitable as candidate genes for the associations found for fumonisin resistance.

GRMZM2G060216 (176 Mbp, bin 3.06), GRMZM2G083347 (14 Mbp, bin 10.03) and GRMZM2G104789 (5.4 Mbp, bin 5.01) genes were described as involved in response to jasmonic acid (JA) (Supplementary Table 2). The signaling pathway of JA promotes downstream activation of defense genes responsive at PR (pathogenesis-related) proteins, such as chitinases (Lanubile et al., 2012). Hormonal signalizing via salicylic acid, auxin, abscisic acid, ethylene, and by own jasmonic acid, are orchestrated until they reach the nucleus (Berens et al., 2017; Lanubile et al., 2014; Wang et al., 2016). Besides this, the GRMZM2G060216 gene refers to the transcription factor LG2, which is related to systemic acquired resistance by the salicylic acid-mediated signaling pathway (Chen et al., 2012).

The candidate gene GRMZM2G135013 (233 Mbp, bin 2.09 - Table 2), described as a G-type lectin S-receptor-like serine, contains a receptor kinase protein proposed as plant sensors at pathogen invasion. Previous studies showed that lectin receptor kinases are further divided into three types (G-, C-, and L-type) based on the extracellular lectin motifs (Wang et al., 2017). In soybean, this protein is identified as a putative abiotic stress-responsive gene (Sun et al., 2012) in a microarray study (Ge et al., 2010).

GRMZM2G036708 gene was described as a cysteine synthase (bin 2.06, 107 Mbp). The largest class of resistance proteins involved in ETI response consists of nucleotide-binding-leucine rich repeat (NB-LRR) proteins (Samayoa et al., 2019). The cysteine synthase that like other several PRRs, has leucine-rich receptor-like kinases, that were also identified in studies of associated with Fusarium ear rots are one of the greatest challenges for maize consumption chain. This disease is resistance in maize (Lanubile et al., 2014; Wang et al., 2016).

Coan et al. (2018) also reported SPN significantly associated with Fusarium ear rot in bin 10.03, in SNP physical position 234 Mbp. Wisser et al. (2006) found bin 10.03

to contain a large QTL conditioning resistance to several maize diseases. Therefore is important for resistance since common rust resistance genes rp1 and rp5 were found in this bin (Chen et al., 2016; Coan et al., 2018). Already the gene GRMZM2G104789 (54 Mbp, bin 5.01- Table 2) was annotated as a serine/threonine-protein kinase HT1. This sequence is involved in the control of stomatal movement in response to CO2 (Hashimoto et al., 2006) and is also described to be involved in plant defense responses (Poland et al., 2011, Wang et al., 2012). Several SNPs associated with the candidate genes with protein domains that have high similarity to the pathogenesis-related proteins and were reported to enhance disease resistance.

The gene GRMZM2G154156 (107 Mbp, bin 2.05- Table 2) was described for protein Ubiquitin ligase complex. The regulatory process in ubiquitylation is specific resides in at least two proteins, the E3 ligase and the cognate substrate, and a number of abiotic stresses that are mediated by protein ubiquitylation processes (Haak et al., 2017). Members from this family are involved in the regulation of some biological processes, including vegetative growth, plant reproduction, biotic and abiotic stresses tolerance. Another gene GRMZM2G022213 (208 Mbp, bin 2.08 - Table 2) to annotated for zinc finger protein MAGPIE that usually regulates tissue boundaries cell division and asymmetric cell division (Welch et al., 2007).

GRMZM2G014618 and GRMZM2G051270 genes located in the bin 5.01, 7 Mbp, to correspond a sulfate adenylyltransferase cysteine (Table 2). ATP-S could be involved in plant-tolerance to several abiotic stresses via different S-compounds pathogen responses (Álvarez et al., 2012). S-containing compounds directly or indirectly modulated/regulated by ATP-S are involved in plant tolerance to both biotic and abiotic stresses (Anjum et al., 2015). There is a high correlation between fumonisin contamination, linoleic acid content and masking action in maize hybrids with higher

oleic to linoleic ratio (Dall'Asta et al., 2012). This masking phenomenon consists of the formation of covalent bonds between the tricarballylic groups of fumonisins and sulfhydryl groups of the side chains of amino acids in proteins. The genes GRMZM2G014618 and GRMZM2G051270 present the sulfate groups and might be related to the increase of fatty acid composition on fumonisin contamination and the occurrence of hidden fumonisins in maize.

The SNPs linked to candidate genes significantly associated with fumonisin resistance could be used as molecular markers decrease contamination to this mycotoxin. The unknown genes or not directly involved genes have the potential to be detailed and might be involved in resistance, since biochemical and genetic pathways leading to resistance to fumonisin are complex and, for the most part, unknown (Zila et al., 2014). Thus, the SNPs associated with fumonisin resistance in this study will be useful for whole-genome selection in tropical maize.

The complex nature of resistance has made it difficult for maize breeders to effectively incorporate novel resistance alleles into adapted breeding pools; as a result, most commercial maize hybrids have lower levels of resistance than desired (Bush et al., 2004). Therefore, to used inbred lines that presented a higher frequency of favorable alleles and lower fumonisin content could be used in future crosses for the generation of resistant hybrids, supporting advances in plant breeding.

7 CONCLUSIONS

Fumonisin presented substantial additive polygenic variation, thus phenotypic and genomic selection approaches should be effective, as long as quality phenotypic resistance evaluations can be performed to allow direct selection or provide training data for genomic selection models. We found 45 significantly associated SNPs and, nineteen QTLs for fumonisin, which co-localized with previously reported linkage QTL, while others were novel, that have the potential to be used as molecular markers for fumonisin resistance.

8 REFERENCES

Álvarez, C., Bermúdez,M.A., Romero,L.C., Gotor,C., and García, I. (2012). Cysteine homeostasis plays na essential role in plantimmu- nity. *New Phytol.* 193, 165–177.doi:10.1111/j.1469-8137.2011. 03889.x.

Andorf, C. M., C. J. Lawrence, L. C. Harper, M. L. Schaeffer, D. A. Campbell et al., (2010). The Locus Lookup tool at MaizeGDB: identification of genomic regions in maize by integrating sequence information with physical and genetic maps. Bioinformatics 26: 434–436.

Anjum, N. A., Gill, R., Kaushik, M., Hasanuzzaman, M., Pereira, E., Ahmad, I., Gill, S. S. (2015). ATP-sulfurylase, sulfur-compounds, and plant stress tolerance. *Frontiers in plant science*, *6*, 210. doi:10.3389/fpls.2015.00210

Berens, M. L., Berry, H. M., Mine, A., Arguesco, C. T., and Tsuda, K. (2017). Evolution of hormone signaling networks in plant defense. Annu. Rev.Phytopathol. 55, 401–425. doi: 10.1146/annurev-phyto-080516-035544

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S.L. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics Applications Note*, v. 23, p. 2633–2635.

Bolduan, C., T. Miedaner, W. Schipprack, B.S. Dhillon, A.E. Melchinger: 2009. Genetic variation for resistance to ear rots and mycotoxin contamination in earlyEuropean maize inbred lines. Crop Sci. 49:2019–2028.

Boutrot, F., and Zipfel, C. (2017) Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. Annu. Rev. Phytopathol. 55, 257–286. doi: 10.1146/annurev-phyto-080614-120106.

Browning, B.L., S.R. Browning. (2016). Genotype imputation with millions of reference samples. The American Journal of Human Genetics 98(1):116–126

Battilani, P., Pietri, A., Barbano, C., Scandolara, A., Bertuzzi, T., and Marocco, A. (2008). Logistic regression modeling of cropping systems to predict fumonisin contamination in maize. J. Agric. Food Chem. 56, 10433–10438. doi: 10.1021/ jf801809d.

Bush, B.J., M.L. Carson, M.A. Cubeta, W.M. Hagler, G.A. Payne. 2004. Infection and fumonisin production by Fusarium verticillioides in developing maize kernels. Phytopatholog. 94:88–93.

Butler, D.G., B.R. Cullis, A.R. Gilmour, and B.J. Gogel. 2009. ASReml-R reference manual. Release 3. Tech. Rep. Queensland Dep. Primary Ind., Brisbane, QLD.

Cheng, S.-F., Y.-P. Huang, Z.-R. Wu, C.-C. Hu, Y.-H. Hsu et al. (2010) Identification of differentially expressed genes induced by bamboo mosaic virus infection in Nicotiana benthamiana by cDNA-amplified fragment length polymorphism. BMC Plant Biol. 10: 286.

Chern, M., Bai, W., Sze-To, W.H., Canlas, P.E., Bartley, L.E., and Ronald, P.C. (2012). A rice transient assay system identifies a novel domain in NRR required for interaction with NH1/ OsNPR1 and inhibition of NH1-mediated transcriptional activation. Plant Methods. 8, 6.

Conesa, A., S. Götz, J. M. García-Gómez, J. Terol, M. Talón et al., (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674– 3676.

Cullis, B., P. Jefferson, R. Thompson, and A.B. Smith. (2014) Factor analytic and reduced animal models for the investigation of additive genotype-by-environment interaction in outcrossing plant species with application to a *Pinus radiata* breeding program. Theor. Appl. Genet. 127:2193–2210. doi:10.1007/s00122-014-2373-0.

Chaudhary, H.K., Kaila V., Rather S.A. (2014) Maize. In: Pratap A, Kumar J (eds) Alien gene transfer in crop plants: achievemnets and impacts. Springer, New York, pp 27–51

Chen, J., Shrestha R., Ding J. et al (2016) Genome-wide association study and QTL mapping reveal genomic loci associated with Fusarium ear rot resistance in tropical

maize germplasm. G3:genes. Genomes Genet 6:3803–3815. https ://doi. org/10.1534/g3.116.03456 1

Chilaka, A.C., De Boevre M., Atanda O.O,. De Saeger S. (2016) Occurrence of Fusarium mycotoxins in cereal crops and processed products (Ogi) from Nigeria. Toxins (Basel) 8:1–18. https :// doi.org/10.3390/toxin s8110 342

Coan, M.M.D., Senhorinho, H.J.C., Pinto R.J.B et al (2018) Genome-wide association study of resistance to ear rot by *Fusarium verticillioides* in a tropical field maize and popcorn core collection. Crop Sci 58:564–578. https ://doi.org/10.2135/crops ci201 7.05.0322

Clements, M. J., Maragos, C. M., Pataky, J. K., and White, D. G. (2004). Sources of resistance to fumonisin accumulation in grain and Fusarium ear and kernel rot of corn. Phytopathology 94, 251–260. doi: 10.1094/PHYTO.2004.94.3.251

Endelman, J.B., and J.K. Jannink. (2012). Shrinkage estimation of the realized relationship matrix. G3 (Bethesda) 2:1405–1413. doi:10.1534/g3.112.004259

Eller, M.S., J.B. Holland, G.A. (2008). Payne. Breeding for improved resistance to fumonisin contamination in maize. Toxin Rev. 27(3–4):371–89.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, S.E. Mitchell. (2011). A robust, simple genotyping-bysequencing (GBS) approach for high diversity species. *PLoS One*. 6:1–10.

Ferrigo, D., Raiola, A., and Causin, R. (2016). Fusarium toxins in cereals: occurrence, legislation, factors promoting the appearance and their management. Molecules 21, 1–35. doi: 10.3390/molecules21050627

Gaikpa, D.S. & Miedaner, T. Theor Appl Genet (2019). Springer-Verlag GmbH Germany, part of Springer Nature. 1432-2242. doi:10.1007/s00122-019-03412-2.

Ge, Y., Li, Y., Zhu, Y.-M., Bai, X., Lv, D.-K., Guo, D., et al. (2010). Global transcriptome profiling of wild soybean (Glycine soja) roots under NaHCO3 treatment. BMC Plant Biol. 10:153. doi: 10.1186/1471-2229- 10-153

Gelderblom, W. C. A., Jaskiewicz, J., Marasas, W. F. O., Thiel, P. G., Horak, R. M., Vleggar, R., et al. (1988). Fumonisins novel mycotoxins with cancer promoting activity produced by Fusarium moniliforme. Appl. Environ. Microbiol. 54, 1806–1811.

Gelderblom, W. C. A., Snyman, S. D., Abel, S., Lebepe-Mazur, S., Smuts, C. M., Van der Westhuizen, L., et al. (1996). “Hepatotoxicity and carcinogenicity of the fumonisins in rats. A review regarding mechanistic implications for establishing

risk in humans," in Fumonisins in Food, eds L. S. Jackson, J. W. De Vries, and L. B. Bullerman (New York, NY: Plenum Press), 279–296. doi: 10.1007/978-1- 4899-1379-1_24

Gilmour, A.R., Thompson, R., and Cullis, B.R. (1995). AI, an efficient algorithm for REML estimation in linear mixed models. Biometrics 51, 1440–1450.

Giomi, G.M., Kreff, E.D., Iglesias J et al (2016) Quantitative trait loci for Fusarium and Gibberella ear rot resistance in Argentinian maize germplasm. Euphytica 211:287. https ://doi.org/10.1007/s1068 1-016-1725-z

Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire., Q. Sun, et al. 2014. A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*. 9:903–916.

Haak, D. C., Fukao, T., Grene, R., Hua, Z., Ivanov, R., Perrella, G., & Li, S. (2017). Multilevel Regulation of Abiotic Stress Responses in Plants. *Frontiers in plant science*, *8*, 1564. doi:10.3389/fpls.2017.01564

Harper, L.C., M.L. Schaeffer, J. Thistle, J.M. Gardiner, C.M. Andorf, D.A. Campbell, et al. 2011. The MaizeGDB genome browser tutorial: One example of database outreach to biologists via video. Database 2011:bar016. doi:10.1093/database/bar016

Hashimoto, M., Negi J., Young J., Israelsson M., Schroeder J.I., Iba K. 2006. *Arabidopsis* HT1 kinase controls stomatal movements in response to CO_2. *Nature Cell Biology* **8**: 391– 397.

Lanubile, A., Pasini, L., Lo Pinto, M., Battilani, P., Prandini, A., and Marocco, A. (2011). Evaluation of broad spectrum sources of resistance to *Fusarium verticillioides* and advanced maize breeding lines. World Mycotoxin J. 1, 43–51. doi: 10.3920/WMJ2010.1206.

Lanubile, A., Ferrarini, A., Maschietto, V., Delledonne, M., Marocco, A., and Bellin, D. (2014). Functional genomic analysis of constitutive and inducible defense responses to *Fusarium verticillioides* infection in maize genotypes with contrasting ear rot resistance. BMC Genomics 15:710. doi: 10.1186/1471-2164- 15-710

Lanubile A., Maschietto V., Borrelli V.M. et al (2017) Molecular basis of resistance to Fusarium ear rot in maize. Front Plant Sci 8:1–13. https ://doi.org/10.3389/fpls.2017.01774.

Li H., R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 25:1754–1760.

Logrieco, A., Mulè, G., Moretti, A., and Bottalico, A. (2002). Toxigenic Fusarium species and mycotoxins associated with maize ear rot in Europe. Eur. J. Plant Pathol. 108, 597–609. doi: 10.1023/A:1020679029993.

Löffler M., T. Miedaner, B. Kessel, M. Ouzunova. 2010. Mycotoxin accumulation and corresponding ear rot rating in three maturity groups of European maize inoculated by two Fusarium species. Euphytica. 2010;174(2):153–64.

Maschietto, V, Colombi C, Pirona R et al (2017) QTL mapping and candidate genes for resistance to Fusarium ear rot and fumonisin contamination in maize. BMC Plant Biol 17:1–21. https ://doi.org/10.1186/s1287 0-017-0970-1

Mesterházy, Á., Lemmens M, Reid L.M. (2012) Breeding for resistance to ear rots caused by *Fusarium* spp. in maize - A review. Plant Breed 131:1–19. https ://doi.org/10.1111/j.1439-0523.2011.01936 .x.

Missmer, S. A., Suarez, L., Felkner, M., Wang, E., Merrill, A. H. Jr., Rothman, Munkvold GP, Desjardins AE (1997) Fumonisins in maize: Can we reduce their occurrence? Plant Dis 81:556–565. https ://doi.org/10.1094/PDIS.1997.81.6.556

Poland, J.A., P.J. Bradbury, E.S. Buckler, and R.J. Nelson. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. Proc. Natl. Acad. Sci. USA 108:6893–6898. doi:10.1073/pnas.1010894108

RCore Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Robertson-Hoyt, L.A., M.P. Jines, P.J. Balint-Kurti, C.E. Kleinschmidt, D.G. White,G.A. Payne, C.M. Maragos, T.L. Molnar, J.B. Holland. 2006. QTL mapping for Fusarium ear rot and fumonisin contamination resistance in two maize populations. Crop Sci. 46(4):1734–43.

Romay, M.C., M.J. Millard, J.C. Glaubitz, J.A. Peiffer, K.L. Swarts, T.M. Casstevens, et al. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14:R55. doi:10.1186/gb-2013-14-6-r55

Schnable, P.S., Ware D., Fulton R.S., Stein J.C., Wei F., Pasternak S., Liang C., Zhang J., Fulton L., Graves T.A., Minx P., Reily A.D., Courtney L., Kruchowski S.S., Tomlinson C., Strong C., Delehaunty K., Fronick C., Courtney B., Rock S.M., Belter E., Du F., Kim K., Abbott R.M., Cotton M., Levy A., Marchetto P., Ochoa K., Jackson S.M., Gillam B. et al. 2009The B73 maize genome: complexity, diversity, and dynamics. Science. 326:1112–1115.

Santiago, R., Cao A. 2015. Butron A. Genetic factors involved in fumonisin accumulation in maize kernels and their implications in maize agronomic management and breeding. Toxins. 7(8):3267–96.

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen, R.W. Allard. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *ProcNatlAcadSci*.USA 81:8014–8018.

Stacklies, W., H. Redestig, M. Scholz, Walther D. and Selbig J. 2007. pcaMethods – a Bioconductor package providing PCA methods for incomplete data. Bioinformatics, 23, 1164-116.

Sun, X., R. Elston, N. Morris, and X. Zhu. 2013. What is the significance of difference in phenotypic variability across SNP genotypes? Am. J. Hum. Genet. 93:390–397. doi:10.1016/j.ajhg.2013.06.017.

Wang, Y, Zhou Z, Gao J et al (2016) The mechanisms of maize resistance to *Fusarium verticillioides* by comprehensive analysis of RNA-seq data. Front

Wang, Y, Bouwmeester K (2017) L-type lectin receptor kinases: New forces in plant immunity. PLoS Pathog. 2017;13(8):1164.

Welch, D., Hassan, H., Blilou, I., Immink, R., Heidstra, R. and Scheres, B. (2007). Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORTROOT action. Genes Dev. 21, 2196-2204.

Wisser, R. J., Balint-Kurti, P. J., and Nelson, R. J. (2006). The genetic architecture of disease resistance in maize: a synthesis of published studies. Phytopathology 96, 120–129. doi: 10.1094/PHYTO-96-0120.

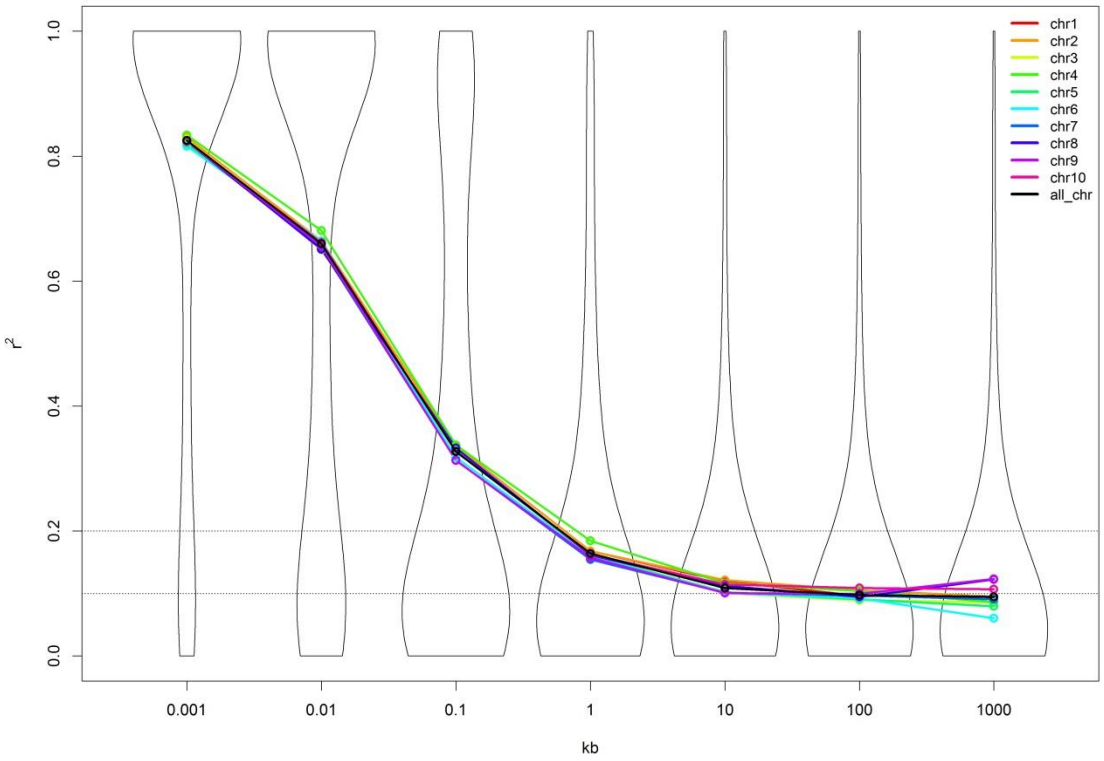
Yan, J., M. Warburton, and J. Crouch. 2011. Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. Crop Sci. 51:433. doi:10.2135/cropsci2010.04.0233.

Zhang, X., Valdés-López, O., Arellano, C., Stacey, G., and Balint-Kurti, P. (2017). Genetic dissection of the maize (Zea mays L.) MAMP response. Theor. Appl. Genet. 130, 1155–1168. doi: 10.1007/s00122-017-2876-6.

Zila, C.T., Samayoa L.F., Santiago R. et al (2013) A genome-wide association study reveals genes associated with Fusarium ear rot resistance in a maize core diversity panel. G3: genes. Genom Genet 3:2095–2104. https ://doi.org/10.1534/g3.113.00732 8.

Zila, C.T., Ogut F., Romay M.C. et al (2014) Genome-wide association study of Fusarium ear rot disease in the U.S.A. Maize inbred line collection. BMC Plant Biol 14:1–15. https ://doi.org/10.1186/ s1287 0-014-0372-6.

9 SUPPLEMENTAL MATERIAL



Supplementary Figure 1 - Genome-wide average linkage disequilibrium (LD) decay over genetic distances. Plot of pair-wised single-nucleotide polymorphism (SNP) LD r^2 values as a function of intermarker map distance (kb). Violin Plot pair distance and mean chromosome (colored lines) and general (black line) imbalances in the maize panel.

Supplementary Table 1- Selection of 10% susceptible and 10% resistant lines in maize panel according to associative mapping and other summary statistics for the SNPs significantly associated with fumonisin resistance SNP associated.

	10% resistent	10% susceptible
Mean lines	0.543	6.881
Mean favorable alleles	0.573	0.253
Mean favorable homozygotes	0.538	0.217
Mean Favorable heterozygote	0.072	0.073
Mean of unfavorable homozygotes	0.391	0.710
Selection Differential	-3.157	3.181
Select gain	-0.853	1.776

Supplementary Table 2 - Genes and description according to GO terms.

GO	Number	Proteína
response to cadmium ion	7	Zm00001d013296_P001, Zm00001d004413_P001, Zm00001d004413_P002, Zm00001d018331_P004, Zm00001d013164_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
negative regulation of transcription, DNA-templated	5	Zm00001d042777_P001, GRMZM2G083347_P02, Zm00001d042777_P002, AC210013.4_FGP013, Zm00001d013164_P001
response to cytokinin	5	Zm00001d013296_P001, Zm00001d004413_P001, Zm00001d004413_P002, GRMZM2G051270_P02, GRMZM2G051270_P03
response to wounding	4	Zm00001d004413_P001, Zm00001d004413_P002, Zm00001d018331_P004, Zm00001d013164_P001
positive regulation of transcription, DNA-templated	4	Zm00001d042777_P001, Zm00001d042777_P002, Zm00001d013164_P001, Zm00001d006682_P001
protein localization	4	Zm00001d018331_P004, GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
response to jasmonic acid	4	Zm00001d042777_P001, GRMZM2G083347_P02, Zm00001d042777_P002, Zm00001d013164_P001
protein ubiquitination	3	AC210013.4_FGP013, Zm00001d018331_P004, Zm00001d004407_P003
Golgi vesicle transport	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
bone development	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
selenium compound metabolic process	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
response to cold	3	Zm00001d042777_P001, Zm00001d042777_P002, AC210013.4_FGP013
phosphorylation	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
negative regulation of cellular process	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
organic substance transport	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
regulation of cellular biosynthetic process	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
response to auxin	3	Zm00001d042777_P001, Zm00001d042777_P002, Zm00001d013164_P001
regulation of lipid biosynthetic process	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
3'-phosphoadenosine 5'-phosphosulfate biosynthetic process	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
sulfate assimilation	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
nitrogen compound transport	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
protein trimerization	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
hydrogen sulfide biosynthetic process	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03
response to molecule of bacterial origin	3	Zm00001d042777_P001, GRMZM2G135013_P02, Zm00001d042777_P002
anther development	3	Zm00001d042777_P001, GRMZM2G083347_P02, Zm00001d042777_P002
ion transport	3	GRMZM2G013200_P02, GRMZM2G013200_P01, GRMZM2G013200_P03
cellular response to sulfate starvation	3	Zm00001d013296_P001, GRMZM2G051270_P02, GRMZM2G051270_P03

GO	Number	Proteína
response to ethylene	2	Zm00001d013288_P001, Zm00001d013164_P001
cysteine biosynthetic process from serine	2	Zm00001d004413_P001, Zm00001d004413_P002
leaf senescence	2	GRMZM2G083347_P02, Zm00001d006682_P001
hydrogen peroxide mediated signaling pathway	2	Zm00001d042777_P001, Zm00001d042777_P002
positive regulation of leaf development	2	Zm00001d042777_P001, Zm00001d042777_P002
response to salt stress	2	GRMZM2G083347_P02, Zm00001d013164_P001
defense response to bacterium	2	Zm00001d042777_P001, Zm00001d042777_P002
response to freezing	2	GRMZM2G135013_P02, Zm00001d013164_P001
detoxification of nitrogen compound	2	Zm00001d004413_P001, Zm00001d004413_P002
fruit ripening, climacteric	2	Zm00001d004413_P001, Zm00001d004413_P002
cyanide catabolic process	2	Zm00001d004413_P001, Zm00001d004413_P002
pollen tube growth	2	Zm00001d004413_P001, Zm00001d004413_P002
immune response	2	Zm00001d004413_P001, Zm00001d004413_P002
response to xenobiotic stimulus	2	Zm00001d042777_P001, Zm00001d042777_P002
systemic acquired resistance, salicylic acid mediated signaling pathway	2	Zm00001d042777_P001, Zm00001d042777_P002
protein phosphorylation	2	Zm00001d013288_P001, Zm00001d018331_P004
pollen development	2	GRMZM2G083347_P02, Zm00001d013164_P001
root hair cell development	2	Zm00001d004413_P001, Zm00001d004413_P002
response to abscisic acid	2	GRMZM2G083347_P02, Zm00001d013164_P001
asymmetric cell division	1	Zm00001d006682_P001
response to insect	1	GRMZM2G135013_P02
cellular component assembly	1	Zm00001d004407_P003
seed morphogenesis	1	GRMZM2G083347_P02
protein-containing complex subunit organization	1	Zm00001d004407_P003
blue light signaling pathway	1	Zm00001d013164_P001
response to flooding	1	GRMZM2G083347_P02
regulation of reproductive fruiting body development	1	GRMZM2G083347_P02
response to karrikin	1	Zm00001d049616_P001
response to organic cyclic compound	1	Zm00001d013022_P001
ion homeostasis	1	Zm00001d013164_P001
protein localization to nucleus	1	Zm00001d006682_P001
photosynthesis	1	AC210013.4_FGP013
regulation of meristem growth	1	Zm00001d006682_P001
regulation of cutin biosynthetic process	1	Zm00001d013164_P001
response to oxygen-containing compound	1	Zm00001d013022_P001
regulation of cell division	1	Zm00001d006682_P001
regulation of raffinose metabolic process	1	Zm00001d013164_P001
endodermal cell differentiation	1	Zm00001d013164_P001
abscisic acid homeostasis	1	GRMZM2G083347_P02
negative regulation of transcription by RNA polymerase II	1	Zm00001d004407_P003
positive regulation of apoptotic signaling pathway	1	Zm00001d004407_P003
photoperiodism, flowering	1	Zm00001d006682_P001
regulation of auxin polar transport	1	Zm00001d006682_P001
suberin biosynthetic process	1	Zm00001d049616_P001

GO	Number	Proteína
xylem development	1	GRMZM2G083347_P02
response to carbohydrate	1	Zm00001d013288_P001
anther wall tapetum development	1	Zm00001d013164_P001
cellular response to organic substance	1	Zm00001d013288_P001
cutin biosynthetic process	1	Zm00001d049616_P001
plant-type cell wall organization	1	Zm00001d013164_P001
chaperone-mediated protein complex assembly	1	Zm00001d018331_P004
cuticle pattern formation	1	Zm00001d013164_P001
negative regulation of cell population proliferation	1	Zm00001d013164_P001
heat acclimation	1	Zm00001d018331_P004
negative gravitropism	1	Zm00001d006682_P001
cellular biosynthetic process	1	Zm00001d013288_P001
positive regulation of protein sumoylation	1	Zm00001d004407_P003
circumnutation	1	Zm00001d006682_P001
cellular response to interleukin-6	1	Zm00001d004407_P003
positive regulation of biosynthetic process	1	GRMZM2G013200_P03
cellular response to interleukin-7	1	Zm00001d018331_P004
regulation of epidermal cell differentiation	1	Zm00001d006682_P001
regulation of root morphogenesis	1	Zm00001d013164_P001
chloroplast organization	1	Zm00001d013288_P001
flower development	1	Zm00001d049616_P001
oxidation-reduction process	1	Zm00001d013022_P001
multidimensional cell growth	1	Zm00001d013164_P001
trichome branching	1	Zm00001d013164_P001
regulation of developmental process	1	Zm00001d013288_P001
regulation of embryonic development	1	GRMZM2G083347_P02
mitotic cell cycle process	1	Zm00001d004407_P003
seed maturation	1	Zm00001d006682_P001
organic substance biosynthetic process	1	Zm00001d013288_P001
fatty acid metabolic process	1	Zm00001d013022_P001
vegetative to reproductive phase transition of meristem	1	AC210013.4_FGP013
leaf morphogenesis	1	Zm00001d006682_P001
regulation of seed germination	1	Zm00001d006682_P001
morphine biosynthetic process	1	Zm00001d013022_P001
response to UV	1	Zm00001d013164_P001
anatomical structure development	1	Zm00001d013288_P001
pollen sperm cell differentiation	1	Zm00001d049616_P001
response to high light intensity	1	Zm00001d018331_P004
terpene metabolic process	1	Zm00001d013022_P001
axillary shoot meristem initiation	1	Zm00001d013164_P001
regulation of secondary cell wall biogenesis	1	Zm00001d013164_P001
sulfate reduction	1	GRMZM2G051270_P03
fruit ripening	1	GRMZM2G083347_P02
Group II intron splicing	1	AC210013.4_FGP013
detection of gravity	1	Zm00001d006682_P001

GO	Number	Proteína
protein autophosphorylation	1	GRMZM2G135013_P02
floral organ morphogenesis	1	Zm00001d006682_P001
determination of adult lifespan	1	Zm00001d018331_P004
positive gravitropism	1	Zm00001d006682_P001
response to absence of light	1	GRMZM2G083347_P02
cellular response to inorganic substance	1	Zm00001d013288_P001
cellular response to heat	1	Zm00001d018331_P004
positive regulation of ethylene biosynthetic process	1	GRMZM2G083347_P02
microtubule-based process	1	Zm00001d004407_P001
menthol biosynthetic process	1	Zm00001d013022_P001
positive regulation of auxin biosynthetic process	1	Zm00001d006682_P001
integument development	1	GRMZM2G083347_P02
intracellular signal transduction	1	Zm00001d013288_P001
cellular response to cold	1	Zm00001d018331_P004
embryo development ending in seed dormancy	1	GRMZM2G083347_P02
regulation of timing of transition from vegetative to reproductive phase	1	Zm00001d006682_P001
lignin biosynthetic process	1	Zm00001d013164_P001
response to stress	1	Zm00001d013288_P001
actin cytoskeleton organization	1	Zm00001d004407_P003
small molecule metabolic process	1	Zm00001d013288_P001
plant-type hypersensitive response	1	Zm00001d042777_P001
response to unfolded protein	1	Zm00001d018331_P004
regulation of protein localization	1	Zm00001d004407_P003
negative regulation of signal transduction	1	Zm00001d013288_P001
response to lipid	1	Zm00001d013022_P001
activation of MAPKKK activity	1	Zm00001d004407_P003
regulation of ERK1 and ERK2 cascade	1	Zm00001d004407_P003
positive regulation of gibberellic acid mediated signaling pathway	1	Zm00001d006682_P001
organic substance catabolic process	1	Zm00001d013022_P001
regulation of starch metabolic process	1	Zm00001d006682_P001
response to gibberellin	1	Zm00001d013164_P001
regulation of small molecule metabolic process	1	GRMZM2G013200_P03
sieve element differentiation	1	GRMZM2G083347_P02
negative regulation of ATPase activity	1	Zm00001d018331_P004
response to abiotic stimulus	1	Zm00001d013288_P001
response to salicylic acid	1	Zm00001d013164_P001
response to chitin	1	Zm00001d013164_P001
gibberellic acid homeostasis	1	Zm00001d006682_P001
cellular response to oxygen-containing compound	1	Zm00001d013288_P001
positive regulation of cell differentiation	1	Zm00001d013164_P001
regulation of stomatal movement	1	Zm00001d013288_P001
plant-type cell wall modification involved in multidimensional cell growth	1	GRMZM2G083347_P02
response to hydrogen peroxide	1	Zm00001d018331_P004
reproduction	1	Zm00001d018331_P004

Supplementary Table 3 - Regions and genes that have already been identified according to associative and QTLs mapping for resistant ear rot or fumonisin.

Reference	Chrom	Bin	Position start	Position End	Gene candidate	Function
Maschietto et al, 2017	2		45192080	48148553	GRMZM2G031331	Barley mlo defense gene homolog3 R Modulated
					GRMZM2G104843	Lipoxygenase 8 RTS Modulated
					GRMZM2G331701	22.0 kDa class IV heat shock protein RTS Modulated
					GRMZM2G086971	Thioredoxin RTS Constitutive
					GRMZM2G053111	Serine threonine-protein kinase-like ccr4 ST Modulated
	7		50.591.357	130000000	GRMZM2G461159	S-adenosylmethionine decarboxylase MP Modulated
					GRMZM2G477743	Monoglyceride lipase-like MP Modulated
					GRMZM2G093947	RNA-binding protein rbp37 MP Constitutive
					GRMZM2G133613	Avr9 elicitor response protein R Modulated
					GRMZM2G328877	Heat shock protein 70 RTS Constitutive
					GRMZM2G139535	Heat shock factor-transcription factor 21 RTS Modulated
					GRMZM5G813217	Heat shock protein 83-like RTS Modulated
					GRMZM2G153607	Early-responsive to dehydration stress-related protein RTS Modulat
					GRMZM2G139815	WRKY 74 transcription factor ST Modulated
					GRMZM2G334165	Cysteine-rich receptor-like protein kinase 10 ST Modulated
					GRMZM2G025761	Transcriptional adaptor family protein ST Constitutive
					GRMZM2G123119	APETALA2/ethylene responsive element binding protein transcription
					GRMZM2G467943	APETALA2/ethylene responsive element binding protein transcription
					GRMZM2G141219	APETALA2/ethylene-responsive transcription factor at1g16060-like S
					GRMZM2G052667	APETALA2/ethylene responsive element binding protein transcription
					GRMZM2G092137	OCS element-binding factor 1 ST Modulated
					GRMZM2G009045	Phosphate carrier mitochondrial-like T Modulated
					GRMZM2G075951	Carbohydrate transmembrane transporter T Constitutive
					AC234166.1_FG	002 Hypothetical protein ZEAMMB73_317354

Reference	Chrom	Bin	Position start	Position End	Gene candidate	Function
Zila et al. 2013	1		63540590	-	GRMZM2G703598	
	5		30997717		GRMZM2G111477	
	9		151295233		GRMZM2G178880	
Zila et al. 2014	5		64771372		GRMZM2G060659	
	9		19532465		GRMZM2G035665	
	4		7566354		GRMZM2G372364	
	4		7618125		GRMZM2G012821	
	4		9353851		GRMZM2G419836	
	4		124930006		GRMZM2G106752	
Chen et al. 2012		1.0	2786055		GRMZM2G041881	3UTRNascentpolypeptideassociated
		1.0	10506267		GRMZM2G028469	Promoter—
		1.1	285314047		GRMZM2G110295	3UTRAntifreezeprotein
		1.1	286228712		GRMZM2G178341	3UTRRibosomalproteinS13
		2.0	8733661		GRMZM2G443445	ExonGroES-like
		2.0	53583850		GRMZM2G069093	PromoterPlantperoxidase
		3.0	10791638		GRMZM2G024551	3UTR—
		3.1	128563291		GRMZM2G175968	Promoter—
		3.1	187947934		GRMZM2G085392	ExonDensegranuleGra7protein
		3.1	201056001		AC207628.4	IntronIQcalmodulin-bindingregion
		4.0	1497071		GRMZM2G156346	PromoterFlagellarmotorswitchprotein
		4.1	183999530		GRMZM2G115499	Exon—
		4.1	217656184		GRMZM2G702806	Exon2-oxoglutarate(2OG)and
		4.1	217656207		GRMZM2G702806	Exon2-oxoglutarate(2OG)and
		4.1	217656309		GRMZM2G702806	Exon2-oxoglutarate(2OG)and
		5.0	72324287		GRMZM2G029879	IntronCyclin-related
		5.0	172983404		GRMZM2G128146	PromoterGlucose/ribitoldehydrogenase glycerol acyltransferase family protein
		5.0	172990198		GRMZM2G128228	Exon—
		6.1	121834796		GRMZM2G341027	Exon—

Reference	Chrom	Bin	Position start	Position End	Gene candidate	Function
		9.0	11972127		GRMZM2G467169	3UTR
		9.0	37162489		GRMZM2G034318	Promoter—
		9.0	85677755		GRMZM2G095206	ExonGlucose/ribitoldehydrogenase
		9.0	85678508		GRMZM2G095206	5UTRGlucose/ribitoldehydrogenase
		9.1	150241000		GRMZM2G148057	IntronKinaseinteracting(KIP1-like) famly protein
		10.0	11675413		GRMZM2G413943	Exon—
		10.0	30829449		GRMZM2G010669	5UTRTranscriptionfactor,MADS-box
		10.0	30829471		GRMZM2G010669	5UTRTranscriptionfactor,MADS-box
		10.0	31526825		GRMZM2G560307	Promoter—
		10.0	32154695		GRMZM2G544512	Promoter—
		10.0	32493898		GRMZM2G027431	5UTRPutativeendonucleaseor glycosyl hidrolase
		10.0	32979981		GRMZM2G109783	Promoter Protein kinase C
Ju et al. 2016	1		6433914		GRMZM2G008122	(AHA3,ATAHA3,HA3)H(+)-ATPase3Stressresponserelated
	1		9580795		GRMZM2G415390	Diseaseresistance-responsivefamilyproteinDiseaseresistancerelated
	1		9807793		GRMZM2G009818	Leucine-richrepeattransmembraneproteinkinasefamilyproteinDiseaseresistancerelated
	1		9811052			
	1		40812382		GRMZM5G891990	(ATPUB13,PUB13)plantU-box13Diseaseresistancerelated
	1		50100218		GRMZM2G394212	AD/NAD(P)-bindingoxidoreductasefamilyproteinDiseaseresistancerelated
	1		64239355		GRMZM2G147698	(ATMYB61,MYB61)mybdomainproteinStressresponserelated
	1		244559899		GRMZM2G120085	Subtilasefamilyproteinother
	2		14836550		GRMZM2G303118	P-loopcontainingnucleosidetriphosphatehydrolasessuperfamilyEnzymesrelatedto
	2		184967793		GRMZM2G414252	(HEC1)basichelix-loop-helix(bHLH)DNA-bindingsuperfamily
	2		197597371		GRMZM2G117865	S-locuslectinproteinkinasefamilyproteinDiseaseresistancerelated
	2		197598070			
	2		202176728		GRMZM2G154864	Transducin/WD40repeat-likesuperfamilyprotein
	2		202178253			Resistance related
	2		202178298			

Reference	Chrom	Bin	Position start	Position End	Gene candidate	Function
	2		221753785		GRMZM2G106560	(ATWRKY75,WRKY75)WRKYDNA-bindingprotein75Diseaseresistancerelated
	3		145476172		GRMZM2G397948	ubiquitin-proteinligasesDiseaseresistancerelated
	3		147411182		GRMZM2G138342	(ATNADK-1,NADK1)NADkinase1
	3		147411924		Disease	resistancerelated
	3		187896997		AC213654.3_	Stressresponserelated
	4		4867002		GRMZM2G468260	RING/U-boxsuperfamilyproteinRegulationprocess
	4		4867006			
	4		198337279		GRMZM2G071405	Pentatricopeptiderepeat(PPR-like)superfamilyproteinother
	4		198337821			
	4		198560823		AC234156.1	
	5		16096144		GRMZM2G160619	Leucine-richrepeatproteinkinasefamilyproteinDiseaseresistancerelated
	5		16844527		GRMZM2G077828	(ATCNGC7,)cyclicnucleotidegatedchannel7
	5		16844540			Disease resistance related
	5		16845040			
	5		16845047			
	5		56372680		GRMZM2G099255	Core-2/I-branchingbeta-1,6-N-acetylglucosaminyltransferase
	5		182033002		GRMZM2G300771	(AGC2,AGC2-1,AtOXI1,OXI1)AGC(cAMP-dependentcGMPdependent kinase family protein
	5		195185908		GRMZM2G154628	aquaporinproteinputativeexpressedStressresponserelated
	7		107926924		GRMZM2G019183	trehalosesynthaseputativeexpressedDiseaseresistancerelated
	7		131442441		GRMZM2G018044	EukaryoticaspartylproteasefamilyproteinEnzymesrelatedto hydrolysis
	7		158749101		GRMZM2G333980	(ATPGIP1,PGIP1)polygalacturonaseinhibitingproteinDiseaseresistancerelated
	7		165360093		GRMZM2G111224	inter-alpha-trypsininhibitorheavychain-relatedother
	7		170031566		GRMZM2G055607	P-loopcontainingnucleosidetriphosphatehydrolasessuperfamilyEnzymesrelatedto hydrolysi
	7		174747321		GRMZM2G434792	ThiaminepyrophosphatedependentpyruvatedecarboxylasefamilyRegulationprocess
	8		13790236		GRMZM2G176568	Homeodomain-liketranscriptionalregulatorRegulationprocess
	8		13790237			
	9		153449080		GRMZM2G374986	Homeodomain-likesuperfamilyproteinRegulationprocess

Reference	Chrom	Bin	Position start	Position End	Gene candidate	Function
Coan et al.,2018	7	7.0	167625331		GRMZM2G459841	Polypeptide: GRMZM2G459841_P01
	10	10.0	38588712		GRMZM2G459820	Polypeptide: full-length complementary
	5	10.0	1870260		GRMZM2G032680	Protein-coding polypeptide
		5.1			GRMZM2G010095	–
	5	5.1	187023602		GRMZM2G153359	Gibberellin 2-oxidase4 (ga2ox4)
					GRMZM2G153356	Polypeptide: GRMZM2G153356_T01
					GRMZM2G131245	Polypeptide: ubiquitin-protein ligase/zinc
					GRMZM2G131254	Enzyme: Ras-related protein RHN1//heterotrimeric G-protein GTPase//
	2	2.1	234935151		GRMZM2G440932	
					AC207888.3_FG009	
	2	2.0	9620998		GRMZM2G388051	
	3	3.0	39211277		GRMZM2G024733	
	7	7.0	157050973		GRMZM2G060993	Enzyme: cis-zeatin O-b-D-glucosyltransferase//flavonol 3-O-glucosyltransferase//flavonol 7-O--glucosyltransfera
	7	7.1	168786006		GRMZM2G039757	Polypeptide: full-length complementary DNA clone ZM_BFc0042D11