

ASSOCIAÇÃO DE MODELO LINEAR MISTO E MAPEAMENTO DE
HERDABILIDADE GENÔMICA REGIONAL PARA IDENTIFICAÇÃO
DE LOCOS E GENES DE CARACTERÍSTICAS RELACIONADAS A
CRESCIMENTO E PRODUTIVIDADE EM MILHO-PIPOCA

GABRIELLE SOUSA MAFRA

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO – UENF

CAMPOS DOS GOYTACAZES – RJ
FEVEREIRO – 2020

ASSOCIAÇÃO DE MODELO LINEAR MISTO E MAPEAMENTO DE
HERDABILIDADE GENÔMICA REGIONAL PARA IDENTIFICAÇÃO
DE LOCOS E GENES DE CARACTERÍSTICAS RELACIONADAS A
CRESCIMENTO E PRODUTIVIDADE EM MILHO-PIPOCA

GABRIELLE SOUSA MAFRA

“Tese apresentada ao Centro de Ciências e
Tecnologias Agropecuárias da Universidade
Estadual do Norte Fluminense Darcy Ribeiro,
como parte das exigências para obtenção do
título de Doutora em Genética e Melhoramento
de Plantas.”

Orientador: Prof. Antônio Teixeira do Amaral Junior

CAMPOS DOS GOYTACAZES – RJ
FEVEREIRO – 2020

FICHA CATALOGRÁFICA

UENF - Bibliotecas

Elaborada com os dados fornecidos pela autora.

M187

Mafra, Gabrielle Sousa.

Associação de modelo linear misto e mapeamento de herdabilidade genômica regional para identificação de locos e genes de características relacionadas a crescimento e produtividade em milho-pipoca / Gabrielle Sousa Mafra. - Campos dos Goytacazes, RJ, 2020.

114 f. : il.

Bibliografia: 68 - 102.

Tese (Doutorado em Genética e Melhoramento de Plantas) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciências e Tecnologias Agropecuárias, 2020.
Orientador: Antonio Teixeira do Amaral Junior.

1. MLMA. 2. RHM. 3. Genômica. 4. Milho-pipoca. I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. II. Título.

CDD - 631.5233

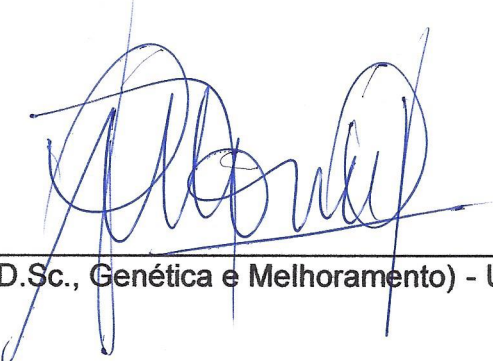
ASSOCIAÇÃO DE MODELO LINEAR MISTO E MAPEAMENTO DE
HERDABILIDADE GENÔMICA REGIONAL PARA IDENTIFICAÇÃO
DE LOCOS E GENES DE CARACTERÍSTICAS RELACIONADAS A
CRESCIMENTO E PRODUTIVIDADE EM MILHO-PIPOCA

GABRIELLE SOUSA MAFRA

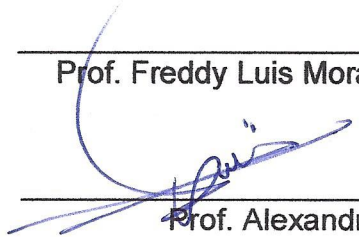
“Tese apresentada ao Centro de Ciências e
Tecnologias Agropecuárias da Universidade
Estadual do Norte Fluminense Darcy Ribeiro,
como parte das exigências para obtenção do
título de Doutora em Genética e Melhoramento
de Plantas.”

Aprovada em 19 de fevereiro de 2020.

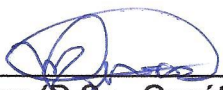
Comissão Examinadora:



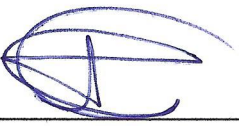
Prof. Freddy Luis Mora Poblete (D.Sc., Genética e Melhoramento) - UTALCA



Prof. Alexandre Pio Viana (D.Sc., Produção Vegetal) - UENF



Prof. Marcelo Vivas (D.Sc., Genética e Melhoramento de Plantas) - UENF



Prof. Antônio Teixeira do Amaral Junior (D.Sc., Genética e Melhoramento) - UENF
(Orientador)

A Deus,

OFEREÇO

*Aos meus pais, Valneide (in memoriam) e José,
Ao meu irmão, Jarley,
E à minha dinda, Ana Lúcia,
Pelo amor e apoio,*

DEDICO

AGRADECIMENTOS

A Deus, por ser minha força e meu sustento.

À minha mãe, Valneide Sousa (*in memoriam*), minha maior incentivadora, ao meu pai José Mafra, por me amar em detalhes, e ao meu irmão Iarley Mafra, meu companheiro.

Aos meus familiares, pelo apoio incondicional nas horas difíceis e nos momentos de alegria.

À Universidade Estadual do Norte Fluminense Darcy Ribeiro e ao Programa de Pós-graduação em Genética e Melhoramento de Plantas, pela oportunidade de aprendizado e crescimento profissional.

À Universidad de Talca - Chile e ao Instituto de Ciencias Biológicas, pela oportunidade de realização do Doutorado Sanduíche.

À FAPERJ, pela concessão das bolsas de Doutorado e Doutorado Sanduíche, e auxílio financeiro para o desenvolvimento do projeto.

Ao meu orientador, Professor Antônio Teixeira do Amaral Junior, pela oportunidade e pelos ensinamentos.

Ao Professor Freddy Mora, pela orientação durante o período de Doutorado Sanduíche, amizade e auxílio durante os meses no Chile.

Ao Professor Marcelo Vivas, pela coorientação, amizade, apoio e ensinamentos durante os anos de pós-graduação.

Ao Dr. Janeo Eustáquio de Almeida Filho, pelo apoio, incentivo, amizade, paciência e ajuda com as análises estatísticas.

Aos Professores Alexandre Pio Viana e Messias Gonzaga Pereira, pelas valiosas contribuições.

Ao Dr. Carlos Maldonado, pela ajuda com as análises estatísticas.

Aos professores do PPGMP, pelos ensinamentos repassados.

Ao secretário do Programa de Genética e Melhoramento de Plantas, José Daniel, pelos auxílios e conselhos.

Aos companheiros do Grupo de Melhoramento de Milho-pipoca.

Aos amigos conquistados na UENF, Juliana Saltires, Guilherme Pena, Fernanda Rossi, Paulo Ricardo, Yure de Souza, Marcelo Moura, Rafael “Cabral”, Fernando Ferreira, Fernando Higino, José Arantes, Cássio Vittorazzi, Amanda Guimarães, Júlio Vettorazzi, Fábio Tomaz e Elisangela Knoblauch.

Aos amigos, Jaomara, Laíne, Leonara, Silvia, Paloma e Raphael, que foram fundamentais para que a vida se tornasse mais leve.

Ao Sr. Geraldo, pela amizade e orientação no campo.

Aos companheiros da Universidad de Talca, Paulina Ballesta, Camilo Valenzuela e Carlos Maldonado, por toda a ajuda durante os meses de convívio.

Ao Dr. Paulo Albuquerque, por acreditar que eu posso ir além.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

A todos, meu muito obrigada!

SUMÁRIO

RESUMO	vii
ABSTRACT	ix
1. INTRODUÇÃO	1
2. OBJETIVOS	4
3. REVISÃO	5
3.1 Milho-pipoca e o melhoramento genético no Brasil e na UENF	5
3.2 A cultura do milho e a era genômica.....	9
3.3 Detecção de QTL via Associação de Modelo Linear Misto (MLMA)	13
3.4 Mapeamento de Herdabilidade Regional (RHM)	16
4. MATERIAL E MÉTODOS	19
4.1 População de estudo e fenotipagem.....	19
4.2 Genotipagem e controle de qualidade	20
4.3 Desequilíbrio de ligação e análise de estrutura populacional	21
4.4 Análises estatísticas da MLMA	22
4.5 Análises estatísticas do RHM	23
4.6 Identificação de genes candidatos.....	24
5. RESULTADOS	26
5.1 Caracterização da população de estudo.....	26
5.2 Herdabilidades capturadas pela MLMA e médias fenotípicas	29
5.3 Comparação entre herdabilidades genômicas por MLMA e RHM	30
5.4 Associação de Modelo Linear Misto (MLMA).....	33

5.5 Associação via Mapeamento de Herdabilidade Regional	39
5.6 Análise comparativa dos resultados da Associação de Modelo Linear Misto (MLMA) e Mapeamento de Herdabilidade Regional (RHM).....	52
6. DISCUSSÃO	53
6.1 Caracterização da população de estudo.....	54
6.3 Comparação de herdabilidades genômicas da MLMA e do RHM	55
6.3 Associação de Modelo Linear Misto (MLMA).....	58
6.4 Análise comparativa dos resultados da Associação de Modelo Linear Misto (MLMA) e do Mapeamento de Herdabilidade Regional (RHM).....	61
CONCLUSÕES	66
REFERÊNCIAS BIBLIOGRÁFICAS.....	68

RESUMO

MAFRA, Gabrielle Sousa; D.Sc.; Universidade Estadual do Norte Fluminense Darcy Ribeiro; Fevereiro, 2020; Associação de modelo linear misto e mapeamento de herdabilidade genômica regional para identificação de locos e genes de características relacionadas a crescimento e produtividade em milho-pipoca; Orientador: Prof. Dr. Antônio Teixeira do Amaral Junior; Conselheiros: Prof. Dr. Alexandre Pio Viana e Prof. Dr. Marcelo Vivas.

A identificação de genes responsáveis por características complexas tem grande potencial para acelerar o melhoramento genético, porém, em milho-pipoca, estudos dessa ordem ainda são limitados. Dessa forma, nesta pesquisa, foi aplicado o método de associação baseada em modelo linear misto (MLMA, sigla em inglês) e mapeamento de herdabilidade regional (RHM, sigla em inglês) para seis importantes características do milho-pipoca. O trabalho foi realizado na população de polinização aberta UENF-14, oriunda de oito ciclos prévios de seleção recorrente intrapopulacional. O polimorfismo de 200 genitores oriundos de oito ciclos de seleção recorrente da UENF-14 apresentou um painel denso de SNPs. Após a realização dos filtros para exclusão das observações perdidas e frequência alélica inferior a 0,05, o estudo pôde ser realizado com o genótipo de 196 indivíduos e 10.507 SNPs, e, desses indivíduos, em 98, foram obtidas progênies S_1 , nas quais foram avaliadas as características: altura de plantas e de espiga; peso de 100 grãos; capacidade de expansão; rendimento de grãos; e volume de pipoca expandido por hectare. As avaliações fenotípicas foram realizadas em dois locais: Campos dos Goytacazes (RJ) e Itaocara (RJ), em

delineamento de blocos incompletos. Com as observações fenotípicas das progênies S_1 e com a caracterização do genoma dos genitores, foram realizadas MLMA e RHM e, posteriormente, verificados genes catalogados nas plataformas MaizeGDB e Phytozome, que potencialmente estão em desequilíbrio de ligação com SNPs associados às características avaliadas. O estudo permitiu observar que, para a maioria das características, a herdabilidade capturada pela MLMA foi maior que a observada no RHM. Além disso, permitiu revelar genes associados às características em ambas as análises, de forma que, na RHM, houve um maior número de associações identificadas, e somente quatro associações coincidiram entre as análises. Apesar disso, esses resultados podem auxiliar no conhecimento sobre a arquitetura genética dos caracteres avaliados, uma vez que este é o primeiro trabalho na cultura de milho-pipoca, e, assim, permitir que seja realizada a seleção precoce de genótipos com as características desejadas.

ABSTRACT

MAFRA, Gabrielle Sousa; D.Sc.; Universidade Estadual do Norte Fluminense Darcy Ribeiro; February, 2020; Mixed linear model association and regional genomic heritability mapping for loci and gene identification of growth and yield traits in popcorn; Advisor: Prof. Dr. Antônio Teixeira do Amaral Junior; Counselors: Prof. Dr. Alexandre Pio Viana and Prof. Dr. Marcelo Vivas.

The identification of genes responsible for complex characteristics has great potential to accelerate genetic improvement, however, in popcorn, these types of studies are still limited. Thus, in this research, the association method based on mixed linear model (MLMA) and regional heritability mapping (RHM) was applied to six important characteristics of popcorn. The work was carried out in the open-pollination population UENF-14, from eight previous cycles of recurrent intra-population selection. The polymorphism of 200 parents from eight recurrent UENF-14 selection cycles showed a dense panel of SNPs. After performing the filters to exclude lost observations and allele frequency lower than 0.05, the study could be carried out with the genotype of 196 individuals and 10,507 SNPs, and of these individuals, in 98, S1 progenies were obtained, which were evaluated: plant and ear height; 100 grains weight; popping expansion; grain yield; and popping volume per hectare. The phenotypic evaluations were carried out in two locations: Campos dos Goytacazes (RJ) and Itaocara (RJ), in an incomplete block design. With the phenotypic observations of the S1 progenies and with the characterization of the genome of the parents, MLMA and RHM were performed and, subsequently, verified genes cataloged on the MaizeGDB and Phytozome

platforms, which are potentially unbalanced with SNPs associated with the evaluated characteristics. The study allowed us to observe that, for most characteristics, the heritability captured by MLMA was greater than that observed in RHM. In addition, it allowed to reveal genes associated with the characteristics in both analyzes, so that, in RHM, there was a greater number of associations identified, and only four associations coincided between the analyzes. Despite this, can assist in the knowledge about the genetic architecture of the evaluated characters, since this is the first work in the culture of popcorn, and, thus, allow the early selection of genotypes with the desired characteristics

1. INTRODUÇÃO

No milho (*Zea mays* L.), as características mais importantes, economicamente, são geralmente herdadas de forma quantitativa, e a base genética é atribuída a poligenes, bem como a efeitos da interação entre genes e/ou entre genes e o meio ambiente (Xiao et al., 2017). Por ser uma espécie utilizada como modelo pela genética clássica e citogenética, contribui expressivamente para a compreensão de processos fundamentais, que incluem reprodução, fotossíntese, biossíntese de metabólitos primários, elementos móveis e relações estrutura-função cromossômica (Zhou et al., 2009).

Nas últimas décadas, com o avanço das técnicas com marcadores moleculares, a genética quantitativa molecular vem focando em estudos que objetivem inferir sobre a arquitetura genética das características complexas, identificando os *Quantitative Trait Loci* QTLs responsáveis por essas características (Huang e Mackay, 2016). No entanto, dissecá-las tem sido um obstáculo para os melhoristas de plantas e animais, devido ao grande número de genes envolvidos na expressão do caráter e à grande interferência do ambiente no qual são submetidos.

Uma ferramenta utilizada com sucesso para identificar a base genética de características quantitativas é o uso de modelo linear misto em análises de associação (Zhang et al., 2010). O mapeamento de associação refere-se não apenas à identificação de QTLs, mas também à identificação de genes candidatos com base na significância estatística entre marcadores e fenótipo (Paes et al.,

2016). O uso de marcadores moleculares, denominados polimorfismos de nucleotídeo único [*Single Nucleotide Polymorphism* (SNP)], vem contribuindo para o avanço do conhecimento a respeito da estrutura e diversidade genética em populações de milho. A utilização destes marcadores tornou-se fundamental nos estudos de associação de modelo linear misto (Wang et al., 2012; Yang et al., 2014b; Wen et al., 2018).

Desde a liberação do genoma de referência B73 no milho (Schnable et al., 2009), o uso de modelo linear misto tem apresentado forte advento, e dezenas de características importantes para a agricultura foram dissecadas (Xiao et al., 2017). Por ser uma espécie de fecundação cruzada, o milho possui uma diversidade genética abundante e a capacidade de grandes cruzamentos controlados, tornando a espécie um bom modelo para o uso em estudos de associação em seu genoma (Li et al., 2012; Han e Huang, 2013).

O mapeamento de associação foi utilizado com sucesso em milho para identificar regiões genômicas que influenciam a altura da planta (Weng et al., 2011), o desenvolvimento radicular (Pace et al., 2015), a resistência a *Fusarium graminearum* Schwabe (Han et al., 2018) e a *Fusarium verticillioides* (Sacc.) Nirenberg (Coan et al., 2018b), bem como a biossíntese de óleo em grãos (Li et al., 2013). Após o mapeamento por associação em regiões cromossômicas de milho-pipoca, foram encontrados genes candidatos que afetam amido, proteína de armazenamento e conteúdo de óleo de grãos de pipoca e conteúdo de polissacarídeos de pericarpo (Paes et al., 2016). No entanto, pouco se sabe a respeito de genes que estejam envolvidos no controle das principais características da cultura do milho-pipoca, tais como rendimento e capacidade de expansão dos grãos.

Embora o estudo de associação utilizando modelo linear misto seja o método mais popular para analisar a base genética de características quantitativas, para a maioria destas, apenas uma proporção relativamente baixa da variação genética total é localizada (Nagamine et al., 2012; Thorwarth et al., 2018). Isso ocorre porque, provavelmente, os efeitos individuais de alguns SNPs são muito pequenos e, por conseguinte, não são detectados ao serem submetidos a testes de significância rigorosos (Yang et al., 2010).

Uma abordagem analítica, chamada mapeamento por herdabilidade regional, ou simplesmente RHM, foi proposta por Nagamine et al. (2012) com o

objetivo de capturar mais sutilmente a variação genética potencialmente associada a características quantitativas, que não possa ser detectada pela associação genômica básica. Neste método, usa-se uma matriz de parentesco genômico (GRM, sigla em inglês) entre indivíduos, com base em variantes de SNPs comuns e raras encontradas em segmentos curtos do genoma, comumente chamados de “janelas”, para estimar a variância da característica explicada por tais regiões (Shirali et al., 2016; Resende et al., 2017).

Para estimar as herdabilidades regionais e genômicas, utiliza-se um modelo misto baseado na máxima verossimilhança restrita (REML, sigla em inglês) e dois componentes de variância, um atribuído ao genoma inteiro e outro a uma região genômica específica, que são ajustados no modelo (Uemoto et al., 2013).

Pesquisas vêm revelando o maior poder de detecção do RHM, quando comparado ao método de associação genômica ampla (Riggio et al., 2013; Resende et al., 2017). Entretanto, embora o método tenha se difundido nas áreas humana (Shirali et al., 2016; Zeng et al., 2017b) e animal (Matika et al., 2016), são raros os estudos que utilizam esse método em plantas. Até o momento, foram identificados apenas três trabalhos com RHM, desenvolvidos com organismos vegetais, sendo esses: eucalipto (Resende et al., 2017), mandioca (Okeke et al., 2018) e feijão (Resende et al., 2018). Uma abordagem com a metodologia de RHM pode identificar variantes comuns e raras envolvidas na expressão de características complexas, como rendimento e capacidade de expansão de grãos, em populações de milho-pipoca.

Até onde se sabe, esta é a primeira tentativa de usar o procedimento RHM nessa cultura. Devido à importância da população de milho-pipoca UENF-14, um estudo aprofundado a respeito das regiões genômicas responsáveis pelas principais características de interesse pode auxiliar no desenvolvimento de germoplasma de milho-pipoca da UENF. Com isso, os objetivos do trabalho podem ser definidos em:

2. OBJETIVOS

2.1 Geral:

Por meio do estudo de associação via MLM e RHM, identificar QTLs e genes relacionados às principais características de crescimento e produtividade em milho-pipoca.

2.2 Específicos:

2.2.1 Realizar a MLMA e o RHM para seis características de importância agronômica em milho-pipoca, avaliadas em diferentes ambientes;

2.2.2 Identificar potenciais genes, que estejam em desequilíbrio de ligação com SNPs associados.

3. REVISÃO

3.1 Milho-pipoca e o melhoramento genético no Brasil e na UENF

Na cultura do milho-pipoca, é importante considerar que os genótipos melhorados devem atender aos interesses dos produtores e, principalmente, dos consumidores. Características como elevada produtividade de grãos, alta capacidade de expansão, bem como menor altura de plantas e, conseqüentemente, menor altura de inserção da espiga estão entre as mais procuradas, que podem também ser combinadas com resistência ou tolerância a estresses bióticos e abióticos (Li et al., 2007; Gerhardt et al., 2017; Kurosawa et al., 2017a; Oliveira et al., 2018; Amaral Junior et al., 2019; Senhorinho et al., 2019).

No Brasil, o melhoramento de milho-pipoca teve início entre os anos de 1982 e 1984 (Sawazaki, 2010) no Instituto Agronômico de Campinas (IAC). Segundo Sawazaki (2010), no ano de 1941, foi introduzida, no IAC, a variedade "South American Mushroom" (SAM), oriunda dos Estados Unidos da América, mas somente entre os anos de 1946/1947 ela foi lançada. No ano de 1950, foi apresentada a variedade Branca Pontuda. Entre os anos de 1982-1984, teve início o Programa de Melhoramento de milho-pipoca realizado por Sawazaki et al. (1986) e Sawazaki (1996), no qual foram avaliadas seis variedades de milho-pipoca, em cruzamento dialélico, quanto à capacidade de expansão dos grãos. Do ano de 1985 até o ano de 1990, foi realizado o melhoramento das variedades SAM e Guarani via seleção massal; de 1991 a 1993, a seleção recorrente

recíproca nas variedades SAM e IAC-64 (versão amarela da variedade Guarani), e deu-se início à obtenção de linhagens das variedades SAM, Guarani e IAC 64 (Sawazaki et al., 2000). No ano de 1993, o IAC iniciou a avaliação de híbridos norte-americanos em parceria com as empresas Conagra e Weaver (Sawazaki et al., 1998).

Os primeiros registros a respeito do melhoramento de milho comum, via metodologia de milho híbrido, datam de 1932 (Krug et al., 1943), embora, Sawazaki et al. (2000) afirmem que, no ano 2000, sua utilização no melhoramento do milho-pipoca ainda era recente no Brasil. O híbrido simples modificado IAC 112, primeira cultivar registrada no Ministério da Agricultura, Pecuária e Abastecimento (MAPA), data do ano de 1998 (Registro Nacional de Cultivares, 2019), e foi desenvolvida pelo Instituto Agrônomo de Campinas, pioneiro no melhoramento de milho no País.

No ano de 2006, houve o lançamento do híbrido *top-cross* (híbrido simples x variedade) IAC 125, obtido de linhagens da população SAM, de IAC 64 e de geração avançada de um híbrido norte-americano (Sawazaki, 2010). Atualmente, é o único genótipo de milho-pipoca híbrido disponibilizado pelo IAC, uma vez que a cultivar IAC 112 não está sendo produzida e as demais — IAC 268 e IAC 367 — estão em processo de produção de sementes (Instituto Agrônomo de Campinas, 2019).

Instituições públicas como a Universidade Estadual de Maringá – UEM (Scapim et al., 2002; Coan et al., 2018a; Sanches et al., 2019; Senhorinho et al., 2019), a Universidade Federal de Viçosa – UFV (Viana et al., 2014; Mundim et al., 2015; Paes et al., 2016; Almeida et al., 2018b, a; Rahim et al., 2019; Andrade et al., 2019), a Universidade Estadual de Londrina – UEL (Carpentieri-Pípolo et al., 2003, 2005; Hülse de Souza et al., 2012), a Embrapa Milho e Sorgo (Pacheco et al., 1998), o Instituto Agrônomo de Campinas (Instituto Agrônomo de Campinas, 2019), a Universidade Federal do Cariri – UFCA (Lima et al., 2018; Martins et al., 2018) e a Universidade Estadual do Norte Fluminense Darcy Ribeiro – UENF (Amaral Júnior et al., 2013; Vittorazzi et al., 2017; Guimarães et al., 2018b), possuem Banco de Germoplasma e/ou desenvolvem programas de melhoramento de milho-pipoca no Brasil.

No entanto, de acordo com informações disponibilizadas pelo Registro Nacional de Cultivares (2019), as únicas instituições públicas mantenedoras de

registros de cultivares de milho-pipoca, junto ao Ministério da Agricultura, Pecuária e Abastecimento, são o Instituto Agrônômico de Campinas e a Universidade Estadual do Norte Fluminense Darcy Ribeiro.

O programa de melhoramento de milho-pipoca na UENF teve início com a introdução do composto indígena UNB-1, no ano de 1993, levado pelo professor Joachim Friedrich Wilhelm Von Bülow à Instituição. Este composto foi, primeiramente, doado pela Escola Superior de Agricultura Luiz de Queiroz Esalq/USP à Universidade de Brasília (UNB), que, por sua vez, o doou à UENF. Após sua introdução no Banco Ativo de Germoplasma (BAG) da Universidade, o composto foi cruzado com a variedade de milho-pipoca americana, cuja progênie resultante foi cruzada com uma variedade de milho-pipoca que possuía alto rendimento e resistência a *Exserohilum turcicum* (Pass.) K. J. Leonard & E. G. Suggs (helmintosporiose) (Pereira e Amaral Júnior, 2001).

A população resultante foi, então, conduzida a dois ciclos de seleção massal e três retrocruzamentos com a variedade americana, dando origem à UNB-2 (Pereira e Amaral Júnior, 2001), que, por sua vez, foi submetida a dois ciclos de seleção massal, originando a população UNB-2U. Após cinco ciclos de seleção recorrente intrapopulacional, a cultivar UENF-14 foi lançada (Amaral Júnior et al., 2013), e, hoje, encontra-se em 9º ciclo de seleção recorrente (C-9).

Inicialmente, com o objetivo de estimar os componentes genéticos da população UNB-2U e de determinar qual a melhor estratégia para a obtenção de maiores ganhos genéticos para as características rendimento de grãos (RG) e capacidade de expansão (CE), Pereira e Amaral Junior (2001) utilizaram o Delineamento I de Comstock e Robinson (1948), que, segundo os autores, é apropriado para o estudo de estrutura genética de populações. Os resultados desse estudo indicaram que a estratégia de família de irmãos-completos foi a que apresentou as maiores perspectivas de ganhos (9,42% GY e 27,09% CE), quando comparada com a de famílias endogâmicas S_1 (7,93% GY e 19,54% CE); meios-irmãos (6,26% GY e 18,45% CE); seleção massal estratificada (2,33% GY e 12,13% CE); e seleção massal (2,23% GY e 11,77% CE). Dessa forma, os autores recomendaram que o procedimento de seleção recorrente de irmãos-completos fosse adotado no melhoramento intrapopulacional da UNB-2U.

Por conseguinte, o primeiro ciclo (C1) de seleção recorrente da população UNB-2U foi conduzido por meio da utilização de famílias de irmãos-completos, as

quais foram avaliadas no ano agrícola de 1998/1991 no estado do Rio de Janeiro. Foram selecionadas as 30 famílias superiores, de acordo com Falconer (1981), a fim de estimar os ganhos genéticos da população, resultando em um ganho de 10,39% e 4,69% para capacidade de expansão e rendimento de grãos, respectivamente (Daros et al., 2002).

No ano agrícola de 2001/2002, o segundo ciclo (C2) de seleção recorrente foi conduzido. Este, por sua vez, foi realizado com a utilização de famílias endogâmicas (S_1). Um total de 40 famílias foram selecionadas, e os ganhos genéticos foram preditos segundo os critérios estabelecidos pelo índice de Smith (1936) e Hazel (1943), perfazendo um ganho de 17,8% para capacidade de expansão e de 26,95% para rendimento de grãos (Daros et al., 2004).

No terceiro ciclo (C3) de seleção recorrente, a estratégia de famílias de meios-irmãos foi adotada, sendo avaliadas 192 famílias, entre as quais foram selecionadas as 30 melhores de acordo com o índice de seleção proposto por Mulamba e Mock (1978), o que proporcionou ganho de 7,16% para capacidade de expansão e de 10,00% para rendimento de grãos (Santos et al., 2007).

Nos demais ciclos — C4 a C8 —, foram utilizadas famílias de irmãos-completos, e a seleção das famílias superiores ocorreu com base no índice de Mulamba e Mock (1978). Os ganhos obtidos foram: 10,58% para CE e 7,71% para GY no C4 (Freitas Júnior et al., 2009); 6,01% para CE e 8,53% para GY no C5 (Rangel et al., 2011); 10,97% para CE e 15,30% para GY no C6 (Ribeiro et al., 2012); 5,11% para CE e 7,78% para GY no C7 (Freitas et al., 2014); e 3,61% para CE e 4,60% para GY no C8 (Guimarães et al., 2018a).

O programa de melhoramento de milho-pipoca da UENF também possui outras linhas de pesquisa, além da seleção recorrente tradicional (Daros et al., 2004; Amaral Junior et al., 2010; Freitas Júnior et al., 2009; Freitas et al., 2014; Ribeiro et al., 2016; Amaral Junior et al., 2016; Guimarães et al., 2019). São desenvolvidos trabalhos visando à obtenção de genótipos resistentes aos principais patógenos que acometem a cultura, tais como *Puccinia polysora* Underw (Mafrá et al., 2018; Schmitt et al., 2019), *Exserohilum turcicum* (Pass.) K. J. Leonard & E. G. Suggs (Santos et al., 2017b; Kurosawa et al., 2017b), *Bipolaris maydis* (Nisik.) Shoemaker (sinônimo de *Helminthosporium maydis* Nisik. and Myiake) (Kurosawa et al., 2017b; Amaral Júnior et al., 2019) e *Fusarium* spp. (Schwantes et al., 2018) bem como a estresses abióticos causados por solos

deficientes em nitrogênio (Santos et al., 2017a, 2019), fósforo (Gerhardt et al., 2017, 2019; Silva et al., 2019), e estresse hídrico (Kamphorst et al., 2018b, a, 2019; Lima et al., 2019).

No entanto, atualmente, trabalhos vêm sendo desenvolvidos a fim de acelerar o processo de seleção de genótipos superiores nos programas de melhoramento. Nesse sentido, o uso de marcadores moleculares do tipo polimorfismos de nucleotídeo único (SNPs) tem sido uma das principais ferramentas na identificação de marcas associadas às principais características e quantificação de diversidade genética na culturas do milho (Van Inghelandt et al., 2010; Frascaroli et al., 2013; Zhang et al., 2016a; Xu et al., 2017; Mengesha et al., 2017; Ertiro et al., 2017; Luo et al., 2017) e milho-pipoca (Mundim et al., 2015; Paes et al., 2016; Coan et al., 2018a; Senhorinho et al., 2019; Andrade et al., 2019), por ocorrerem com uma frequência muito maior no genoma do que os marcadores SSRs (Van Inghelandt et al., 2010), por exemplo.

O programa de melhoramento genético de milho-pipoca da UENF também vem implementando o uso dessa ferramenta, visando a facilitar a seleção dos genótipos superiores no programa de seleção recorrente (SR) intrapopulacional, de forma que o uso de todos os marcadores disponíveis seja útil para prever o valor genotípico, levando a um maior ganho (Massman et al., 2013). Espera-se que esses venham a auxiliar diretamente na seleção, diminuindo o tempo de cada ciclo de SR. Além disso, almeja-se que os SNPs significativos encontrados venham a contribuir para a seleção das características mais importantes da cultura, por meio da sua utilização em outras linhas de pesquisa desenvolvidas no programa.

3.2 A cultura do milho e a era genômica

O genoma do milho (B73 RefGen_v1) é composto por 85% de sequência repetitiva (Schnable et al., 2009). A quantidade de sequências repetidas é a maior causa de diferença entre os tamanhos dos genomas (Muñoz-Diez et al., 2012; Jian et al., 2017) e podem ser ocasionadas por diferenças de temperatura em distintas localizações geográficas, longitudes e latitudes (Díez et al., 2013). Além disso, tempos de floração e diferenças nos tempos ótimos de floração em *clines* altitudinais afetam indiretamente o tamanho do genoma (Jian et al., 2017; Bilinski et al., 2018).

O tamanho do genoma (GS, sigla em inglês) da espécie *Zea mays* ssp. *mays*, medido pelo coeficiente de variação, varia em até 6% quando comparado a seus parentes silvestres mais próximos (ssp. *parviglumis* e *mexicana*), os teosintes (Díez et al., 2013). Linhas temperadas têm genomas significativamente menores do que as linhas tropicais, que são significativamente menores que o teosinte (Chia et al., 2012), o que ocorre também com as raças cujo GS é menor do que em teosintes (Díez et al., 2013), dados que corroboram a hipótese de que o tamanho do genoma da espécie também pode ser influenciado indiretamente pela seleção natural no desenvolvimento e por características adaptativas (Tenaillon et al., 2016; Bilinski et al., 2018).

Desde o sequenciamento do genoma da linhagem de milho denominada B73, em 2009 (Schnable et al., 2009), pesquisas na área genômica vêm buscando identificar genes e genótipos que sejam importantes comercialmente, de forma a auxiliar no aumento da produtividade das áreas de cultivo.

Com a necessidade de disponibilizar à comunidade científica e/ou acadêmica dados a respeito da cultura do milho, no ano de 2001, surgiu a iniciativa de compilar os dados disponibilizados nas plataformas MaizeDB e ZmDB. Eles dizem respeito a informações sobre o genoma do milho e disponibilização de regiões ESTs, GSSs e sequências de proteínas, entre outras. Entretanto, somente no ano de 2003 essa iniciativa se concretizou, dando origem à plataforma on-line denominada MaizeGDB (Lawrence, 2004), que, atualmente, fornece dados a respeito do genoma de referência B73, entre outros.

O genoma de referência da linhagem B73 possui um tamanho de 5,64 pg / 2C (picogramas – pg – por núcleo 2C), ou aproximadamente 2,3 gigabase, e foi sequenciado, pela primeira vez, no ano de 2009 (Schnable et al., 2009; Díez et al., 2013). A contar de seu lançamento, foi revisado três vezes, utilizando-se a estratégia sequenciamento de BAC (cromossomo artificial bacteriano) (Portwood et al., 2019). No entanto, em 2015, uma nova versão foi disponibilizada, dessa vez, baseando-se no sequenciamento *single-molecule real-time* e mapeamento ótico de alta resolução, que resultou em um aumento de 52 vezes no comprimento do *contig* e progressos na montagem de espaços intergênicos e centrômeros (Jiao et al., 2017).

Muitos são os trabalhos que utilizam o B73 como genoma de referência (Ganal et al., 2011; Chen et al., 2016; Morosini et al., 2017; Zhu et al., 2018; Gage

et al., 2019; Alves et al., 2019; Rossi et al., 2020; Li et al., 2020). No entanto, pesquisas envolvendo análises em caráter genômico em milhos especiais, em particular o milho-pipoca, ainda são incipientes, e, em sua maioria, foram realizados para detecção de genes relacionados às principais doenças que acometem a cultura.

Em estudo de associação genômica ampla em milho e milho-pipoca, para identificar SNPs e genes putativos associados à resistência a cercosporiose causada por *Cercospora zeina* Crous & U. Braun sp. nov., utilizando o genoma B73 como referência, foi verificada a presença de três genes que codificam proteínas relacionadas à via de defesa das plantas, e que podem, após validados, ser usados na seleção assistida por marcadores e auxiliar na compreensão da resistência do milho a essa enfermidade (Kuki et al., 2018).

Em estudo avaliando genótipos de milho e milho-pipoca, visando a identificar regiões genômicas, incluindo genes candidatos associados à resistência à podridão de espiga causada por *Fusarium verticillioides* (Sacc.) Nirenberg, Coan et al. (2018) observaram a presença de quatro genes ligados aos SNPs associados, que codificam proteínas relacionadas à defesa. Anteriormente, avaliando um painel de diversidade de milho, incluindo milho-pipoca, verificaram três marcadores significativamente associados à resistência à podridão de espiga (Zila et al., 2013).

Paes et al. (2016) encontraram associações significativas para o volume de expansão e outras características de qualidade, e genes candidatos envolvidos com lípidos, amido, proteína de armazenamento, polissacáridos e síntese da parede celular, em populações de milho-pipoca oriundas de distintos ciclos de seleção.

Em estudo de associação genômica ampla baseada em haplótipos, avaliou-se ângulo foliar, altura de planta (PH, sigla em inglês), altura de espiga (EH, sigla em inglês) e relação entre EH / PH em linhagens de milho tropical, que envolvem milho comum e milho-pipoca, tendo sido observado um total de 40 blocos de haplótipos significativamente associados às características de interesse (Maldonado et al., 2019).

Mais recentemente, foram encontrados quatro SNPs associados à capacidade de expansão em grãos de milho-pipoca, cujos genes anotados demonstraram estar relacionados a funções associadas ao conteúdo de amido,

que desempenham um papel importante na qualidade da expansão (Senhorinho et al., 2019).

No ano de 2017, Jiao et al. (2017) sugeriram uma nova versão do sequenciamento do genoma de B73, denominada B73 RefGen_v4. Os autores observaram que essa nova versão tem contiguidade 240 vezes maior que o genoma da cultivar de milho PH207, recentemente publicada, e concluíram que novas assembleias podem ser realizadas, de forma a elucidar a base da diversidade fenotípica no milho.

Nesse ensejo, além do B73 (Schnable et al., 2009), outros genomas de milho estão disponíveis na plataforma MaizeGDB. São eles: B104 (Maize Genetics and Genome Database, 2019), CML247 (Lu et al., 2015), Mo17 (Yang et al., 2017; Sun et al., 2018), PH207 (Hirsch et al., 2016) e W22 (Springer et al., 2018). No entanto, ainda são poucos os trabalhos que mencionam a utilização desses genomas como referência.

O genoma B104 é 93% semelhante ao B73, por serem derivados da mesma população. A diferença é que o B104 é facilmente transformável, e sua disponibilidade permite que os pesquisadores trabalhem com seus próprios transformantes (Manchanda et al., 2016).

Devido à sua elevada resistência a doenças, a linhagem CML247 apresenta uma valiosa contribuição para o futuro do melhoramento do milho (Lu et al., 2015). Após o sequenciamento, foi utilizada para verificar se os homólogos não alélicos que demonstraram estar presentes nos genomas de B73 e PH207 também estariam presentes no genoma CML247 (Brohammer et al., 2018). Além disso, foi utilizada na comparação entre os transcritos descritos no genoma da linhagem, e transcritos obtidos nos genomas de B73, PH207 e PHJ89, a fim de identificar se esses representavam alelos divergentes da linhagem (Gage et al., 2019).

O sequenciamento do genoma PH207 foi realizado para complementar e compará-lo com a sequência de referência B73 (Hirsch et al., 2016). Ao utilizarem o genoma de PH207 e B73 para identificação sistemática e comparação de genes de repetição pentatricopeptídeo (PPR, sigla em inglês) entre as duas linhagens, foi identificada a presença de 491 e 456 genes PPR nos genomas B73 e PH207, respectivamente (Chen et al., 2018).

O mapa de haplótipos derivados do genoma Mo17 (Yang et al., 2017; Sun et al., 2018) foi utilizado por Yang et al. (2019) para detectar variantes estruturais polimórficas em linhagens de milho, bem como por Anderson et al. (2019), que desenvolveram uma abordagem para definir inserções de elementos transponíveis variáveis e compartilhadas em conjuntos de genomas, e aplicaram esse método a quatro genomas de milho, entre eles, Mo17 e W22.

Dada a grande diversidade de milho, os genomas de referência disponíveis, derivados de variedades de origem temperadas — que, de acordo com Yang et al. (2019), é o caso dos genomas Mo17 (Sun et al., 2018), W22 (Springer et al., 2018) e B73 (Schnable et al., 2009; Jiao et al., 2017) —, são insuficientes para a caracterização de pan-genoma (Yang et al., 2019). No intento de disponibilizar mais genomas, com maior diversidade genética, foi sugerido um novo sequenciamento e montagem de um genoma de referência, dessa vez, utilizando linhagens de milho tropical (Yang et al., 2019). Os autores sugeriram, ainda, a construção de um pan-genoma baseado não apenas no genoma de milho, mas também de seu ancestral, o teosinte.

Observa-se que há um crescente número de pesquisas sendo desenvolvidas, a fim de elucidar cada vez mais o genoma complexo do milho. A quantidade de linhagens sequenciadas e que possuem dados genômicos dispostos em bancos de dados, como MaizeGDB, vem em auxílio desse entendimento. Espera-se que, no futuro, novas metodologias de sequenciamento sejam disponibilizadas, a fim de completar a lacuna ainda existente no conhecimento a respeito dos genes e funções gênicas da espécie *Zea mays* L.

3.3 Detecção de QTL via Associação de Modelo Linear Misto (MLMA)

Nos organismos animais e vegetais, muitas características são herdadas quantitativamente e influenciadas pelo ambiente e por muitos genes (Zeng, 1994; Schäfer-Pregl et al., 1998). As regiões dentro do genoma que possuem genes associados com uma característica quantitativa são denominadas locos de característica quantitativa (QTLs) (Geldermann, 1975; Collard et al., 2005).

No entanto, não é possível realizar a identificação desses locais via fenotipagem convencional. O advento dos marcadores moleculares, nos anos 1980, veio como uma ferramenta importante para auxiliar nessa identificação (Collard et al., 2005), possibilitando, assim, mapear regiões cromossômicas que

afetam esses caracteres (Toledo et al., 2008). O uso de marcadores moleculares auxilia na seleção, e pode acelerar os ganhos genéticos, levando ao desenvolvimento mais rápido de cultivares. Se marcadores vinculados a QTLs para características de interesse forem identificados, a seleção desses locos poderá ser feita ainda em estágio de muda e, portanto, em um estágio muito anterior à seleção baseada em avaliações fenotípicas (Serba et al., 2015).

Na atualidade, muitos tipos de marcadores vêm sendo utilizados na identificação de QTLs para estudos de genomas de espécies de interesse agrônomico (Pandey et al., 2014; Mora et al., 2015; Pootakham et al., 2015) ou de diversidade natural (Rocha et al., 2014; Alves et al., 2016). No entanto, a utilização de marcadores do tipo polimorfismo de nucleotídeo único (SNP) vem ganhando destaque, por apresentarem eficiência, facilidade de automação. (Guan et al., 2019), abundância e acessibilidade ao *screening* de alto rendimento (Simko et al., 2006).

Das abordagens estatísticas que utilizam marcadores SNPs na dissecação do genoma de espécies vegetais e animais, o uso de modelo linear misto (MLM) está entre as que mais recebem atenção, uma vez que permite a avaliação de vários de níveis de parentesco simultaneamente (Yu et al., 2006), e possibilita realizar a análise na presença de estrutura de população, tanto relacionada à localização geográfica (Kang et al., 2010) quanto ao parentesco familiar (Zhang et al., 2010) ou relação enigmática (*cryptic relatedness*) (Price et al., 2010). Nesse método, a estrutura populacional (Pritchard et al., 2000) é ajustada como efeito fixo, enquanto o parentesco entre os indivíduos é inserido como a estrutura de variância-covariância do efeito aleatório para os indivíduos (Zhang et al., 2010).

O método MLM abrange a construção de uma *genetic relationship matrix* (GRM), e atua modelando a estrutura da amostra, estimando a contribuição da GRM para a variação fenotípica usando um modelo de efeitos aleatórios (com ou sem efeitos fixos adicionais) e computando estatísticas de associações que respondem por este componente da variação fenotípica. Conforme relatado, a principal vantagem da associação de modelo linear misto (MLMA) inclui evitar a detecção de associações falso-positivas, devido à presença de estrutura populacional, aumentando o poder de detecção de associações, quando aplicada uma correção específica para essa estrutura (Yang et al., 2014a).

Para estrutura relacionada à posição geográfica, os marcadores que apresentam maiores frequências alélicas relacionadas à população recebem maior correção. Quando a estrutura está relacionada com o parentesco, há uma contribuição menor dos indivíduos relacionados, a fim de impedir que esses influenciem na análise, devido à estrutura de correlação presente (Yang et al., 2014a).

Korte et al. (2012) verificaram que os SNPs de base apresentaram-se correlacionados, devido à estrutura da população, e concluíram que o modelo misto remove efetivamente essa tendência, independentemente da análise ser feita em um fenótipo por vez, usando um modelo misto de característica única, ou ambos simultaneamente usando modelo misto de várias características.

A abordagem MLMA fornece uma alternativa para detecção de marcadores associados com a característica de interesse em coleções de germoplasma e programas de melhoramento em culturas como milho (Yang et al., 2014b), tabaco (Basirnia et al., 2014) e soja (Zhang et al., 2015a).

Em milho, um estudo utilizando dois grandes painéis de linhagens endogâmicas dos tipos dentado e *flint*, adaptadas às condições agroclimáticas europeias, e empregando a abordagem MLM, permitiu identificar um grande número de QTLs para tolerância ao frio, de forma que foram observados mais QTLs associados à tolerância ao frio nas linhagens do tipo *flint* do que nas do tipo dentado (Revilla et al., 2016).

Wang et al. (2012), comparando as abordagens MLM e GLM (Modelo Linear Generalizado), na identificação de genes candidatos à resistência à *Sphacelotheca reiliana* (Kühn) Clint em linhagens de milho, verificaram que o MLM apresentou resultados melhores, quando comparado com a abordagem GLM, permitindo a identificação de 18 genes candidatos associados à resistência de milho ao carvão do pendão.

Na detecção de QTLs favoráveis e genes candidatos à porcentagem de fibra em algodão, Su et al. (2016), utilizando um modelo linear misto, identificaram 12 locos SNPs associados à característica, entre os quais cinco, que se associaram significativamente à característica, foram detectados com o menor valor *P*.

Outra utilização importante da abordagem MLM é na estimação dos componentes de herdabilidade explicados por marcadores (Zaitlen e Kraft, 2012;

Overby et al., 2015). Essa ferramenta auxilia programas de melhoramento, tanto animal quanto vegetal, pois permite que seja realizada a associação de SNPs e QTLs e a identificação de genes candidatos possivelmente relacionados com características de interesse agrônômicos, e, dessa forma, acelera o processo de obtenção de genótipos melhorados.

3.4 Mapeamento de Herdabilidade Regional (RHM)

O método de identificação de QTL por meio da associação genômica ampla, ainda que muito utilizado, vem sendo bastante criticado, uma vez que, segundo alguns estudos, é pouco eficiente para detectar variação genética total de características complexas (Nagamine et al., 2012; Uemoto et al., 2013).

Nagamine et al. (2012) propuseram um método chamado mapeamento de herdabilidade regional, ou simplesmente RHM. Esta metodologia tem como objetivo principal capturar a maior parte da variância genética associada às características complexas, que não pode ser capturada pela associação genômica. Esse fato pode ocorrer devido à falta do poder da associação genômica ampla em detectar alelos causadores raros, que geram pouca variação associada ao SNP, mas que, em conjunto, podem contribuir com uma grande fração da herdabilidade (Yang et al., 2010; Thornton et al., 2013).

Pesquisas na área humana mostraram que múltiplos locos independentes com diferentes frequências e efeitos alélicos estão comumente situados na mesma região gênica ou em regiões de segmentos curtos (Duerr et al., 2006; Haiman et al., 2007). Esses locos podem não ser detectados por análises de SNPs únicos, uma vez que esse tipo de análise não é tão sensível a ponto de identificar efeitos individuais alélicos relativamente pequenos, mesmo que o efeito cumulativo de todo o loco na variância da característica seja grande (Uemoto et al., 2013).

Diante dessa hipótese, Thornton et al. (2013) realizaram um estudo que visava comparar o poder da associação genômica ampla e de outros métodos para a detecção de variantes genéticas que afetam características complexas, por meio de simulação. Os autores concluíram que o ressequenciamento completo produz apenas uma melhora modesta, quando comparado com o poder dos *chips* tradicionais de SNPs.

Alguns métodos são sugeridos como sendo mais vantajosos que a associação genômica ampla, entre eles estão o SKAT (*Sequence Kernel Association Test*) (Wu et al., 2011) e o RHM (*Regional Heritability Mapping*) (Nagamine et al., 2012; Uemoto et al., 2013). Neste último, um modelo misto baseado na máxima verossimilhança restrita (REML) é usado e dois componentes de variância — um contribuído pelo genoma inteiro e um segundo por uma região genômica específica — são ajustados no modelo para estimar heranças regionais e genômicas, respectivamente. O RHM facilita a captura da variação genética que está associada a cada segmento do genoma, combinando os efeitos de variantes comuns e raras em uma região (Nagamine et al., 2012; Uemoto et al., 2013).

A metodologia de mapeamento de herdabilidade regional ainda não está bem consolidada para espécies vegetais, uma vez que a maior parte das pesquisas utilizando esse tipo de metodologia foi realizada para identificação de QTLs relacionados a características complexas em humanos (Uemoto et al., 2013; Caballero et al., 2015; Shirali et al., 2016; Zeng et al., 2017b).

Resende et al. (2018), ao compararem as metodologias de associação genômica ampla e RHM para características complexas em feijão (*Phaseolus vulgaris* L.), verificaram que a proporção de herdabilidade genômica explicada por RHM foi maior do que a explicada pela associação genômica ampla. Ainda segundo os autores, RHM em espécies autógamas tem potencial de identificar QTL de efeito maior combinando variantes alélicas que poderiam ser efetivamente incorporadas em modelos de previsão de genoma completo e rastreadas através de gerações de genitores usando a seleção assistida por marcadores.

Okeke et al. (2018), estudando uma espécie alógama (*Manihot esculenta* Crantz), também concluíram que o mapeamento de herdabilidade regional é eficaz para entender a arquitetura genética de características complexas desta espécie.

Em eucalipto, Resende et al. (2017) verificaram que a análise por meio de RHM detectou 26 locos de características quantitativas (QTLs) abrangendo 2.191 polimorfismos de nucleotídeo único (SNPs), enquanto a associação genômica ampla detectou apenas 13 associações. Além disso, os QTLs de RHM e associação genômica ampla explicaram individualmente 5 a 15% e 4 a 6% da herdabilidade genômica, respectivamente. Os autores concluíram que o

mapeamento de herdabilidade regional foi superior à associação genômica ampla com relação à captura de maiores proporções de herdabilidade genômica.

Apesar de escassas, pesquisas com espécies vegetais utilizando RHM demonstraram ser promissoras, corroborando o objetivo desse trabalho. Em milho (*Zea mays* L.), até o momento, não foram registrados trabalhos envolvendo essa metodologia, e, em milho-pipoca (*Zea mays* L. var. Everta), estudos dessa natureza são ainda mais raros, uma vez que, apesar de se tratar de um tipo especial de milho, que possui relevância econômica para a indústria alimentícia, é mais negligenciado em pesquisas científicas quando comparado ao milho comum.

4. MATERIAL E MÉTODOS

4.1 População de estudo e fenotipagem

Para o estudo, a população escolhida foi a variedade de polinização aberta, denominada UENF-14 (Amaral Júnior et al., 2013). Esta população é proveniente da população UNB-1, que originou a população-base UNB-2 (Pereira e Amaral Júnior, 2001), e que, após dois ciclos de seleção massal, deu origem à população UNB-2U (Amaral Júnior et al., 2013). Com cinco ciclos de seleção recorrente intrapopulacional, foi lançada a cultivar de polinização aberta UENF-14 (Amaral Júnior et al., 2013), que atualmente encontra-se no nono ciclo (C-9), utilizado neste trabalho.

As progênies S_1 provenientes dessa população foram avaliadas fenotipicamente em dois locais, sendo eles: Escola Técnica Estadual Agrícola Antônio Sarlo, situada a 21°43'14.8"S 41°20'38.3"W, no município de Campos dos Goytacazes (RJ) (ENV1), que possui clima tropical e temperatura média anual de 23,6 °C (Climate-data.org, 2020); e no Campo Experimental da UENF, situado a 21°38'45.2"S 42°03'16.3"W, na Ilha Barra do Pomba, município de Itaocara (RJ), que possui clima tropical e temperatura média anual de 23,0 °C (Climate-data.org, 2020), ambas em agosto de 2016.

Cada experimento foi composto por 98 famílias S_1 . Os ensaios foram delineados em blocos incompletos com três repetições ortogonais. Cada família foi semeada em linha de 5 m de comprimento, com espaçamento entre linhas de 0,90 m e entre plantas de 0,20 m. Em cada cova, foram dispostas três sementes,

na profundidade de 0,05 m, e, aos 21 dias após a emergência, um desbaste foi realizado, de modo a permanecer apenas uma planta por cova. Os tratos culturais foram realizados conforme as recomendações para a cultura.

Em ambos os locais, foram mensuradas as principais características de interesse agrônomo para a cultura, a saber:

1. Altura de planta (PH, sigla em inglês, cm), mensurada em seis plantas competitivas de cada parcela, desde o nível do solo à inserção da folha bandeira, com o auxílio de uma vara graduada;
2. Altura de espiga (EH, sigla em inglês, cm), mensurada em seis plantas competitivas de cada parcela, desde o nível do solo à inserção da primeira espiga, com o auxílio de uma vara graduada;
3. Massa de 100 grãos (100GW, sigla em inglês, g), mensurada com o auxílio de uma balança de precisão;
4. Rendimento de grãos (GY, sigla em inglês, Kg/ha), mensurado após o processo de debulhagem de todas as espigas da parcela, com o auxílio de uma balança digital;
5. Capacidade de expansão dos grãos (PE, sigla em inglês, mL/g). Foram pesados 30 g de grãos de cada parcela, com o auxílio de uma balança de precisão, posteriormente dispostos em sacos de papel do tipo kraft e levados ao micro-ondas durante, aproximadamente, 2 min e 20 seg. Após, os grãos expandidos foram colocados em provetas de 2000 mL, e o resultado dividido pela massa inicial de 30 g de grãos; e
6. Volume de pipoca expandida por hectare (PV, sigla em inglês, m³/ha), que foi obtido por meio da multiplicação entre a produtividade média da parcela e a capacidade de expansão dos grãos ($VP = GY \times PE / 1000$).

4.2 Genotipagem e controle de qualidade

A genotipagem foi realizada em 200 genitores resultantes da seleção recorrente; de 98 foram obtidas progênies S₁ para a avaliação fenotípica, conforme descrito anteriormente. Desses 200 indivíduos, foram recolhidas folhas jovens e realizada a extração de DNA, conforme protocolo descrito por Doyle e Doyle (1990). As amostras de DNA foram enviadas à empresa *Rapid Genomics*, que implementou o método *capture seq* (Neves et al., 2014), usando 5.000 sondas bem distribuídas ao longo do genoma, para obtenção de marcadores

SNPs. Este procedimento gerou 21.442 SNPs, que foram submetidos aos filtros: a) removeram-se indivíduos com mais de 10% de dados perdidos; b) removeram-se SNPs com mais de 5% de dados perdidos; e c) retiraram-se SNPs com *Minor Allele Frequency* – MAF < 0,05; resultando em 10.507 SNPs e 196 indivíduos genotipados.

4.3 Desequilíbrio de ligação e análise de estrutura populacional

Com o painel de 10.507 SNPs e 196 indivíduos, foi calculada a *Genetic Relationship Matrix* (GRM), com o pacote rrBLUP (Endelman, 2011), usando o algoritmo de VanRaden (2008). Esse mesmo painel de dados genotípicos também foi utilizado para a estimativa do desequilíbrio de ligação (LD). O LD foi mensurado por meio da estatística r^2 , que foi calculada entre todos os pares de marcadores pertencentes ao mesmo cromossomo, utilizando-se o software PLINK (Purcell et al., 2007). Para o estudo do decaimento do LD ao longo do genoma, foi ajustado o modelo não linear proposto por Hill e Weir (1988), usando a função *nlm* do software R (R Core Team, 2016).

Com as informações de SNPs dos 196 indivíduos, foi avaliada a presença de subpopulações utilizando o software Structure (Pritchard et al., 2000). O algoritmo pressupõe que os sítios polimórficos estejam em equilíbrio de Hardy-Weinberg e em equilíbrio de ligação (Pritchard et al., 2000); dessa forma, foram realizados dois filtros extras, o primeiro pelo teste exato de equilíbrio de Hardy-Weinberg, removendo SNPs com p-valor < 0,05; posteriormente, foi realizado o LD *pruning* em todos os pares de marcadores de mesmo cromossomo que apresentaram estimativas de r^2 acima de 0,1, tendo sido removido um marcador e mantido o outro. Ambos os filtros foram realizados com o software PLINK, resultando em 739 SNPs.

Os modelos testados consideraram valores de subpopulação (K) variando de um (ausência de subpopulação) até 10. Para cada valor de K, a análise foi repetida 20 vezes de forma independente, sendo a cadeia MCMC configurada com 40.000 iterações, considerando um *burn-in* de 10.000 (30.000 iterações após o *burn-in*). Os resultados do modelo bayesiano de Pritchard et al. (2000) foram avaliados sob o critério de Evanno et al. (2005), que é baseado na estatística delta K. Para a obtenção dos valores de delta K, foi utilizada a plataforma on-line Structure Harvester (Earl e VonHoldt, 2012).

4.4 Análises estatísticas da MLMA

A análise de associação foi realizada para todas as características mensuradas, em cada um dos dois locais avaliados, por meio de *Mixed Linear Model Based Association Analysis* – MLM (Yang et al., 2014a), conforme o seguinte modelo:

$$y = X\beta_1 + SNP_i\beta_i + Z_1b + Z_2u + \varepsilon,$$

em que:

y é o vetor com os fenótipos de uma determinada característica; β_1 é o vetor com efeitos fixos — intercepto, repetição, estrutura populacional e covariáveis como: número de plantas na parcela, contadas logo após o desbaste e umidade de grãos para as características 100GW, GY, PE e PV; β_i é o efeito com o i -ésimo SNP candidato (coeficiente de regressão); b é vetor de efeitos de blocos dentro de repetições; u é o vetor de efeitos poligênicos; e é o vetor de efeitos residuais do modelo; X é a matriz de incidência dos efeitos fixos sistemáticos; SNP_i é o vetor com o número de cópias de um determinado alelo do i -ésimo SNP candidato tomado aleatoriamente como referência, $SNP_i = \{0, 1 \text{ ou } 2\}$; e Z_{1-2} são matrizes de incidência dos efeitos aleatórios.

Foi assumido:

$$b \sim N(0, I_b \sigma_b^2); u \sim N(0, G \sigma_u^2); e \sim N(0, I_n \sigma_e^2)$$

$$cov(u, b') = cov(u, e') = cov(b, e') = 0$$

em que G é a matriz de parentesco genômico [*Genetic Relationship Matrix* (GRM)], calculada com o pacote rrBLUP (Endelman, 2011), seguindo o algoritmo de VanRaden (2008); I_b e I_n são matrizes identidade com ordem igual ao número de blocos incompletos e número de observações, respectivamente; σ_b^2 ; σ_u^2 ; σ_e^2 são os componentes de variância associados a b , u , e e , respectivamente. Esses componentes foram estimados pelo método *Restricted Maximum Likelihood* (REML), usando o algoritmo *Average Information* (AI). O MLM foi ajustado pelo pacote ASReml-R (Butler et al., 2009) na linguagem computacional R 3.2.3 (R Core Team, 2016). Para definir o limiar ideal para declarar uma associação significativa, foi utilizado $\alpha = 0,0001$ ($p \leq 0,0001$, ou $-\log_{10}(p) \geq 4$).

Para obter as médias corrigidas para os efeitos sistemáticos, foi adotado o seguinte modelo:

$$y = X\beta_1 + Z_1b + Z_2p + \varepsilon,$$

em que p é o efeito de progênie, tomado como fixo para estimação das médias ajustadas (LSMeans), enquanto os outros termos são idênticos ao modelo anterior. Esse modelo foi ajustado com o pacote lme4 (Bates et al., 2015), e as médias ajustadas foram obtidas com o pacote lsmeans (Lenth, 2016), ambos disponíveis na linguagem R.

As médias ajustadas foram utilizadas para estimação das correlações fenotípicas entre as características em cada ambiente de forma separada, e de uma característica com ela mesma entre os ambientes. Com essas médias corrigidas, também foi estimada a proporção da variância fenotípica explicada pelos marcadores. Para tanto, foi adotado o seguinte modelo:

$$y_{adj} = \mu + g + e,$$

em que y_{adj} corresponde às médias ajustadas para uma determinada característica, μ é o intercepto do modelo, g é o componente genético explicado pelos SNPs e e é o erro aleatório. Neste modelo, é assumido $g \sim N(0, G\sigma_g^2)$, $e \sim N(0, I\sigma_e^2)$, $cov(g, e') = 0$, sendo que a proporção da variância explicada pelos marcadores (h_m^2) é dada por: $h_m^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$. O ajuste do modelo para estimação de h_m^2 foi obtido pelo software GCTA (Yang et al., 2011), utilizando o método REML com algoritmo AI para estimação dos componentes de variância.

4.5 Análises estatísticas do RHM

O RHM fornece estimativas de herdabilidade para segmentos genômicos contendo efeitos alélicos comuns e raros (Resende et al., 2018). Este modelo de RHM apresentado na equação a seguir foi ajustado usando o pacote Regress (Clifford e McCullagh, 2012), na linguagem computacional R 3.2.3 (R Core Team, 2016).

$$y = X\beta_1 + Z_1b + Z_2r + Z_3u + \varepsilon$$

em que:

y é o vetor com os fenótipos de uma determinada característica; β_1 é o vetor com efeitos fixos — intercepto, repetição, estrutura populacional e covariáveis, como: número de plantas na parcela, contadas logo após o desbaste e umidade de grãos para as características 100GW, GY, PE e PV; b é vetor de efeitos de blocos; r é o vetor de efeitos aditivos genômicos regionais aleatórios; u é o vetor

de efeitos poligênicos; e é o vetor de efeitos residuais do modelo; \mathbf{X} é a matriz de incidência dos efeitos fixos sistemáticos; e $\mathbf{Z}_{1,2,3}$ são matrizes de incidência dos efeitos aleatórios.

Foi assumido:

$$b \sim N(0, I_b \sigma_b^2); r \sim N(0, Greg \sigma_r^2); u \sim N(0, G \sigma_u^2); e \sim N(0, I_n \sigma_e^2)$$

$$cov(u, b') = cov(u, e') = cov(b, e') = 0$$

em que \mathbf{G} é a matriz de parentesco genômico [*Genetic Relationship Matrix* (GRM)], calculada com o pacote rrBLUP (Endelman, 2011), seguindo o algoritmo de VanRaden (2008); \mathbf{Greg} é uma matriz semelhante a \mathbf{G} , mas usando um subconjunto da matriz W . Esses subconjuntos foram determinados por “janelas” genômicas ou “regiões” de 100 kb de comprimento sobrepostos, por 50 kb (por exemplo, as três primeiras regiões são 0-100, 50-150 e 150-200 kb) correspondendo ao LD estimado — ver resultados); I_b e I_n são matrizes identidade com ordem igual ao número de blocos incompletos e número de observações, respectivamente; σ_b^2 ; σ_r^2 ; σ_u^2 ; σ_e^2 são os componentes de variância associados a b , r , u , e e , respectivamente. Esses componentes foram estimados pelo método REML, usando o algoritmo Average Information (AI). O RHM foi ajustado pelo pacote Regress (Clifford e McCullagh, 2012), na linguagem computacional R 3.2.3 (R Core Team, 2016), conforme rotina definida por Resende (2017). Para definir o limiar ideal para declarar uma associação significativa, foi utilizada $\alpha = 0,001$ significativos ($p \leq 0,001$, ou $-\log_{10}(p) \geq 3$). De acordo com Resende et al. (2017), a abordagem RHM requer valores de limiar esperados menores para alcançar o mesmo poder de detecção da abordagem MLMA, provavelmente pelo fato de que os SNPs são tratados como efeitos aleatórios na modelagem RHM, enquanto eles são efeitos fixos em MLMA.

As herdabilidades regionais para cada segmento da janela genômica foram determinadas pela seguinte equação: $h_{RHM}^2 = \sigma_r^2 / (\sigma_r^2 + \sigma_g^2 + \sigma_e^2)$.

4.6 Identificação de genes candidatos

Os QTLs identificados por meio da análise MLMA foram submetidos à análise de identificação de genes utilizando o conjunto de dados públicos do genoma do milho, usando como referência o genoma de B73 (Schnable et al., 2009) versão 3. As anotações funcionais desses genes foram realizadas usando o

navegador do genoma MaizeGDB (Andorf et al., 2010). Os genes que estavam dispostos nos SNPs associados ou adjacentes a eles dentro de uma janela de deslizamento de 100 kb (50 kb à direita e 50 kb à esquerda da posição que se encontra o SNP) foram definidos como genes candidatos.

Na análise de RHM, foi utilizado o navegador Phytozome [<https://phytozome.jgi.doe.gov/pz/portal.html>], porém usando o mesmo genoma de referência (*Zea mays* gene annotation on B73 assembly AGPv3 (Schnable et al., 2009)). As regiões para busca de genes foram definidas de acordo com a posição de início e final de cada uma delas. Os genes que estavam dispostos dentro dessas regiões foram definidos como genes candidatos.

5. RESULTADOS

5.1 Caracterização da população de estudo

Para verificar a cobertura de marcadores, foi obtida a distribuição de frequência dos SNPs ao longo do genoma (Figura 1). Pelos resultados, pode-se observar que o polimorfismo ao longo do genoma foi bem mostrado, e existe boa distribuição dos marcadores polimórficos nos cromossomos.

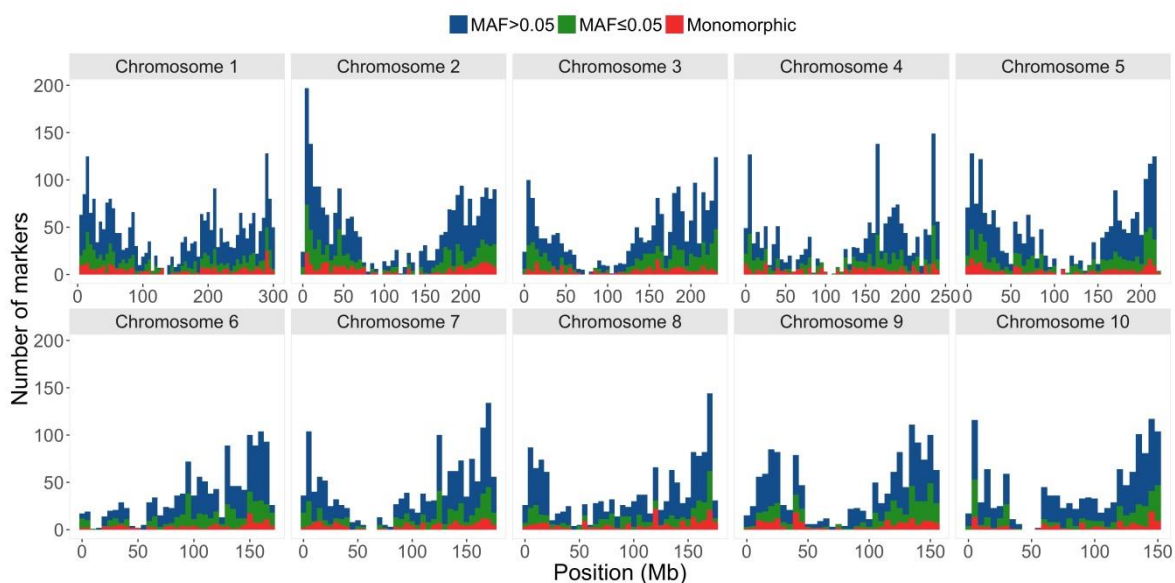


Figura 1. Distribuição de frequência das marcas monomórficas e polimórficas com $MAF \leq 0,05$ e $MAF > 0,05$ ao longo do genoma, considerando janelas de 5 Mb.

Com o intuito de mensurar a magnitude do desequilíbrio de ligação e definir janelas, onde seja esperado que sítios polimórficos estejam em elevado LD, foram estimados o r^2 entre todos os SNPs pertencentes ao mesmo cromossomo. Pelos resultados observados, o maior LD ocorreu no cromossomo 8, cuja metade do decaimento se deu em $\sim 151\text{Kb}$, enquanto o menor LD foi observado no cromossomo 4, onde a metade do decaimento foi observada em $\sim 76\text{Kb}$. Na média dos cromossomos, a metade do LD foi observada em $\sim 110\text{Kb} \pm 6,82$. Em outro estudo de LD com essa mesma população, a metade do LD, considerando todos os SNPs simultâneos, ocorreu em $\sim 107\text{Kb}$ (Schwantes et al., 2020, no prelo).

A fim de evitar associações espúrias, foi avaliada a presença de subpopulações. Com a aplicação do critério de Evanno et al. (2005) sobre os resultados do modelo bayesiano de Pritchard et al. (2000), os resultados indicaram que possivelmente a população está estratificada em três subpopulações (Tabela 1). A Figura 2a indica a probabilidade de cada indivíduo pertencer a cada uma das subpopulações, e, com base no maior valor probabilístico, foi realizada uma classificação dos indivíduos. Assim, as subpopulações 1-3 contemplaram, respectivamente, 30,10; 52,04; e 17,86% dos indivíduos totais; e, respectivamente, 25,51; 54,08; e 20,41% dos indivíduos genitores das progênies S_1 avaliadas.

Para verificar a correspondência da estratificação populacional com o parentesco entre os indivíduos, foi realizada uma análise de componentes principais (PCA, sigla em inglês) com os valores da GRM (Figura 2b). Os resultados da PCA tiveram grande coincidência com os resultados da formação de subpopulações, o que sugere que essa estratificação está relacionada com o parentesco entre os indivíduos, uma vez que estão sob seleção recorrente intrapopulacional.

Contudo, infelizmente, o critério de Evanno et al. (2005) não tem sensibilidade para inferir sobre $K=1$ (ausência de estratificação populacional) e, ao comparar os valores da função de verossimilhança com os modelos assumindo os diferentes valores de K na escala logarítmica ($\log(L)$), o modelo com $K=1$ apresenta ser o mais razoável (Tabela 1). Apesar disso, a fim de evitar associações espúrias, foi utilizado nas análises de MLMA e RHM um $K=2$ (duas subpopulações).

Tabela 1. Resumo dos resultados da análise de estrutura, em que K se refere ao número de subpopulações assumidas no modelo, sendo reportados para cada K a média do logaritmo neperiano (Mean Ln) do máximo da função de verossimilhança (LogL), o desvio padrão (SD) e o Delta K, considerando o método de Evanno et al. (2005)

K*	Mean Ln (K)	SD	Delta K
1	-142684,88	1,55	-
2	-141617,43	14,13	12,64
3	-140728,67	10,75	16,96
4	-140022,2	56,44	2,66
5	-139465,94	50,95	0,13
6	-138903,03	54,62	3,68
7	-138541,11	210,8	0,71
8	-138029,05	246,39	0,74
9	-137700,2	394,68	0,15
10	-137313,46	593,41	-

* Para cada K, foram realizadas 20 repetições de forma independente, totalizando 200 análises. O Delta K não pode ser mensurado para K = 1, nem para o maior valor de K, que, nesse caso, é K = 10

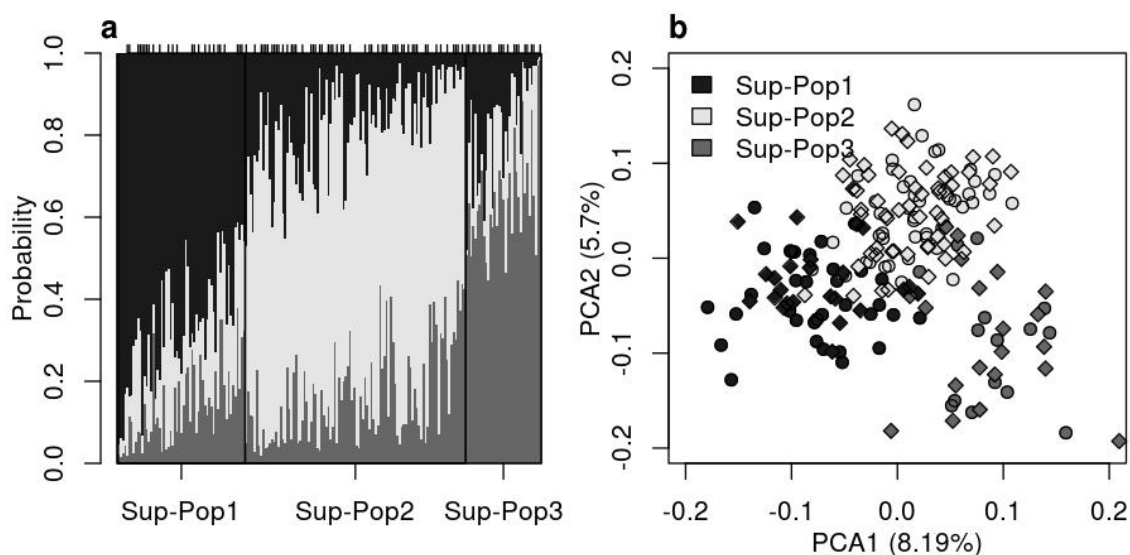


Figura 2. Inferência sobre a estratificação populacional em três subpopulações a partir do modelo Pritchard et al. (2000), sob critério de Evanno et al. (2005). a) Resultados dos valores probabilísticos de cada indivíduo pertencer a determinada subpopulação; b) Análise de Componentes Principais da *Genetic Relationship Matrix*. Em a e b, as classificações das subpopulações foram determinadas pelos valores de maior probabilidade. Marcas na parte superior da figura 2a e pontos com bordas na figura 2b representam os indivíduos genitores das progênies S₁ avaliadas.

5.2 Herdabilidades capturadas pela MLMA e médias fenotípicas

Para as características de maior interesse da cultura GY (2.520,42 Kg/ha) e PE (28,67 mL/g), maiores valores de média foram observados no ENV1, quando comparados com os valores observados no ENV2 (GY = 2.099,94 Kg/ha e PE = 27,75 mL/g). As médias das características EH (84,15 cm) e PH (166,97 cm) no ENV1 foram menores que no ENV2 – EH (113,80 cm) e PH (204,13 cm). Os valores de 100GW foram similares para ambos os locais (13,11 g e 13,41 g, respectivamente), enquanto o valor de PV foi maior no ENV1 (72,55 m³/ha) (Tabela 2).

A proporção dos marcadores que explicam as características fenotípicas (herdabilidade) apresentou valores discrepantes em ambos os locais de avaliação, sendo a característica PH a que apresentou maior valor (\cong 53%) no ENV1, enquanto, no ENV2, a característica que mais se destacou foi PE, com aproximadamente 74% dos SNPs explicando-a, sugerindo que as características responderiam bem à seleção assistida por marcadores SNPs.

Tabela 2. Médias fenotípicas ajustadas e proporção da variância fenotípica média explicada pelos marcadores (h^2m) para seis características avaliadas em milho-pipoca, em Campos dos Goytacazes (ENV1) e Itaocara (ENV2)

Características	ENV1		ENV2	
	Média	h^2m	Média	h^2m
100GW (g)	13,11(0,06*)	0,29(0,30*)	13,41(0,05*)	0,36(0,33*)
EH (cm)	84,15(0,58*)	0,45(0,32*)	113,80(0,95*)	0,11(0,28*)
GY (Kg/ha)	2520,42(42,60*)	<0,01(0,25*)	2099,94(49,12*)	0,37(0,34*)
PE (mL/g)	28,67(0,22*)	0,43(0,31*)	27,75(0,20*)	0,74(0,34*)
PH (cm)	166,97(0,89*)	0,53(0,34*)	204,13(1,24*)	0,37(0,31*)
PV (m ³ /ha)	72,55(1,41*)	0,24(0,31*)	58,65(1,39*)	0,35(0,32*)

* Valores entre parêntesis correspondem ao erro-padrão das estimativas. 100GW = massa de 100 grãos; EH = altura de espiga; GY = rendimento de grãos; PE = capacidade de expansão dos grãos; PH = altura de planta; PV = volume de pipoca expandida por hectare.

As correlações entre as características no ENV1 apresentaram, no geral, valores muito baixos, variando de -0,08 (PH – PE) a 0,9 (PV – GY). A maior

correlação observada neste ambiente ocorreu entre as características volume de pipoca expandida por hectare (PV) e rendimento de grãos (GY). Valores elevados também foram encontrados para a correlação entre altura de planta (PH) e altura de espiga (EH) – 0,82 (Tabela 3), ambos significativos ($P < 0,01$).

No ENV2, os valores variaram de -0,08 (PE – 100GW) a 0,93 (GY – PV). Os maiores valores foram observados entre as características rendimento de grãos (GY) e volume de pipoca expandida por hectare (EH) – 0,93 – e entre altura de espiga (EH) e altura de planta (PH) – 0,83 ($P < 0,01$) (Tabela 3). As características GY e PE, em ambos os ambientes, apresentaram valores de correlação igual a 0,12.

Com relação aos valores de correlação entre as características, considerando as magnitudes, eles variaram de 0,34 (GY) a 0,83 (EH), e todas as associações foram significativas ($P < 0,01$).

Tabela 3. Correlações fenotípicas entre as características avaliadas em cada local de cultivo e entre os dois locais

		ENV1				
ENV2	100GW	EH	GY	PE	PH	PV
100GW	0,71(<0,01)	0,21(0,04)	0,18(0,08)	-0,07(0,47)	0,29(<0,01)	0,13(0,21)
EH	0,19(0,06)	0,83(<0,01)	0,20(0,05)	-0,08(0,41)	0,82(<0,01)	0,13(0,20)
GY	0,24(0,02)	0,23(0,02)	0,34(<0,01)	0,12(0,25)	0,18(0,08)	0,9(<0,01)
PE	-0,08(0,43)	0(0,98)	0,12(0,24)	0,48(<0,01)	-0,08(0,45)	0,52(<0,01)
PH	0,23(0,02)	0,83(<0,01)	0,22(0,03)	0(0,97)	0,8(<0,01)	0,12(0,25)
PV	0,19(0,06)	0,23(0,03)	0,93(<0,01)	0,45(<0,01)	0,22(0,03)	0,31(<0,01)

Diagonal superior: correlação entre as características em Campos dos Goytacazes (ENV1); diagonal inferior: correlação entre as características em Itaocara (ENV2). Diagonal central: correlação entre as características nos dois locais. 100GW = massa de 100 grãos; EH = altura de espiga; GY = rendimento de grãos; PE = capacidade de expansão dos grãos; PH = altura de planta; PV = volume de pipoca expandida por hectare.

5.3 Comparação entre herdabilidades genômicas por MLMA e RHM

Foi utilizado um tamanho de 100 kb, com sobreposição de 50 kb, para definir o tamanho de cada região genômica, cada uma cobrindo um número variável de SNPs polimórficos, que não foram definidos previamente; e, dessa forma, cada região apresentou um número definido de SNPs. Dezenove QTLs

foram mapeados pelo RHM no ENV1, cada um abrangendo entre 2 e 8 SNPs, nos cromossomos 1, 2, 4, 5, 6 e 8 (Tabela 4), enquanto, no ENV2 foram mapeados nove QTLs, abrangendo entre 2 e 9 SNPs, dentro dos cromossomos 2, 4, 6, 7 e 9 (Tabela 5).

Tabela 4. Resultados da detecção de locos de características quantitativas (QTLs) via mapeamento de herdabilidade regional (RHM) usando segmentos genômicos de 0,1-Mb com janela deslizante de 0,05-Mb para seis características quantitativas avaliadas em população de milho-pipoca sob seleção recorrente intrapopulacional, no ambiente de Campos dos Goytacazes, RJ (ENV1)

Característica	Crom.	Posição inicial da região (pb)	Posição final da região (pb)	Número de SNPs na região	h ² REG	-log ₁₀ p
100GW	4	26378603	26478603	2	0,123916	3,922309
100GW	4	26328603	26428603	6	0,188984	3,320643
EH	2	148888849	148988849	4	0,079301	3,17188
EH	2	148838849	148938849	7	0,079618	3,028434
GY	1	48905387	49005387	2	0,127929	3,854726
GY	1	48955387	49055387	2	0,127929	3,854726
GY	1	56805387	56905387	2	0,492078	3,667729
GY	1	56855387	56955387	2	0,492078	3,667729
GY	6	165623194	165723194	5	0,929289	3,467931
GY	2	225088849	225188849	7	0,107897	3,322997
GY	2	225138849	225238849	7	0,107897	3,322997
GY	4	219528603	219628603	2	0,29114	3,266408
GY	4	219578603	219678603	2	0,29114	3,266408
PE	2	13388849	13488849	8	0,560739	4,197226
PE	2	13438849	13538849	8	0,560739	4,197226
PE	5	13560296	13660296	7	0,203652	3,168033
PE	5	13610296	13710296	7	0,203652	3,168033
PH	8	171723438	171823438	3	0,105226	3,084349
PH	8	171773438	171873438	3	0,105226	3,084349

Tabela 5. Resultados da detecção de locos de características quantitativas (QTLs) via mapeamento de herdabilidade regional (RHM) usando segmentos genômicos de 0,1-Mb com janela deslizante de 0,05-Mb para seis características quantitativas avaliadas em população de milho-pipoca sob seleção recorrente intrapopulacional, no ambiente de Itaocara, RJ (ENV2)

Característica	Crom.	Posição inicial da região (pb)	Posição final da região (pb)	Número de SNPs na região	h ² REG	-log ₁₀ p
100GW	6	155323194	155423194	3	0,210555	3,435156
EH	7	115031917	115131917	2	0,119873	3,121412
EH	7	115081917	115181917	2	0,119873	3,121412
EH	4	178478603	178578603	9	0,135972	3,118481
GY	2	200838849	200938849	3	0,078476	3,281193
GY	2	200888849	200988849	3	0,078476	3,281193
PV	9	3856799	3956799	2	0,162408	3,222951
PV	9	3906799	4006799	2	0,162408	3,222951
PV	2	200838849	200938849	3	0,061766	3,098072

No ENV1, QTLs regionais para 100GW, EH e PH estavam localizados em cromossomos individuais, enquanto QTLs para GY e PE foram mapeados em vários cromossomos diferentes (Tabela 4). No ENV2, para as características 100GW e GY, foram identificados QTLs regionais em cromossomos individuais, enquanto, para as demais (EH e PV), foram mapeados em distintos cromossomos. As características GY e PV são correlacionadas, sendo encontrado um QTL em comum entre ambas, no ENV2 – cromossomo 2 (Tabela 5).

As herdabilidades variaram entre 0,079 e 0,92 no ENV 1, enquanto, no ENV2, variaram entre 0,06 e 0,2. Observa-se que estas não foram superiores quanto às verificadas na análise de MLMA (Tabela 2), sendo maior apenas para GY e PE, no ENV1 (Tabela 4), e EH no ENV2 (Tabela 5). No entanto, esse fato pode estar relacionado às várias regiões apresentadas, enquanto, na análise de MLMA, a h² é investigada em todo genoma.

Observa-se que as herdabilidades observadas no método RHM foram inferiores às estimadas pela análise de MLMA (Tabela 2), sendo superiores apenas para GY e PE, no ENV1 (Tabela 4), e para a característica EH no ENV2 (Tabela 5). Frações maiores de herdabilidade foram observadas no RHM para

características relacionadas à produtividade (GY, 100GW e PE), quando comparadas às características de crescimento (EH e PH) no ENV1. Por sua vez, no ENV2, as características responsáveis pelas maiores herdabilidades foram EH, PV e 100GW.

5.4 Associação de Modelo Linear Misto (MLMA)

Para verificar as regiões do genoma associadas com características de interesse em milho-pipoca, foi realizada a análise de associação de modelo linear misto. Na Figura 3, encontram-se os *Manhattan Plot*, evidenciando SNPs significativos ($p \leq 0,0001$, ou $-\log_{10}(p) \geq 4$). A partir dos resultados de LD (Figura 3), foi estabelecida uma "janela" de busca de genes catalogados, que possam estar em elevado LD, com os SNPs associados. Dessa forma, foi padronizada uma janela de 100 Kb, que corresponde a 50 kb à direita e 50kb à esquerda da posição que se encontra o SNP. Nas Tabelas 6 e 7, observam-se com mais detalhes os SNPs associados com as características avaliadas nos ENV1 e ENV2, respectivamente, os quais tiveram genes anotados, por estarem supostamente em elevado LD.

Tabela 6. Genes candidatos obtidos da análise de associação de modelo linear misto para seis caracteres de interesse em milho-pipoca no ambiente de Campos dos Goytacazes, RJ (ENV1)

Nº	Característica	SNP	Crom*	Posição (pb)	MAF ^a	P valor	Anotação
1	EH	GRMZM2G002959	10	124300563	0,255102	0.0000269	Glutaryl-CoA dehydrogenase
2	100GW	GRMZM2G089995	4	26424338	0,1173469	0.0000069	AP2/EREBP-transcription factor 209
3	100GW	GRMZM2G034152	10	62249986	0,2857143	0.0000546	Polyamine oxidase 1
4	GY	GRMZM2G069618	1	56893602	0,0688776	0.0000228	TPR domain containing protein
5	PE	GRMZM2G461936	5	13615937	0,2959184	0.0000889	AGO-108 - argonaute108
6	PH	GRMZM2G118950	2	28937135	0,1173469	0.0000793	AMT-3 - ammonium transporter3
7	PV	GRMZM2G086573	2	4674885	0,1572165	0.0000383	AP2/EREBP-transcription factor 24

* Cromossomo

^a Alelo de menor frequência

Tabela 7. Genes candidatos obtidos da análise de associação de modelo linear misto para seis caracteres de interesse em milho-pipoca avaliados no ambiente de Itaocara, RJ (ENV2)

Nº	Característica	SNP	Crom*	Posição (pb)	MAF ^a	P valor	Anotação
1	EH	GRMZM2G002959	10	124300563	0.255102	0.0000509	Glutaryl-CoA dehydrogenase
2	100GW	GRMZM2G087032	3	161700883	0.0535714	0.0000880	C3H-transcription factor 313
3	PE	GRMZM2G098793	5	58753471	0.1147959	0.0000077	Glycosyltransferase
4	PE	GRMZM2G081048	7	139116617	0.1096939	0.0000022	Oxidoreductase
5	PE	GRMZM2G048672	5	56022372	0.0994898	0.0000339	Macrophage migration inhibitory factor
6	PH	GRMZM2G043435	6	160035981	0.0807292	0.0000875	Respiratory burst oxidase-like protein C
7	PV	GRMZM2G110726	9	3937091	0.2602041	0.0000305	Protein BOBBER 1
8	PV	GRMZM2G020150	6	156592891	0.4540816	0.0000248	AP2/EREBP-transcription factor 196

* Cromossomo

^a Alelo de menor frequência

Comparando os locais físicos de SNPs significativos no genoma de referência B73, os genes candidatos referentes foram identificados. Em Campos dos Goytacazes, os genes apresentaram-se distribuídos nos cromossomos 1, 2, 4, 5 e 10. Para a característica EH, foi identificado o gene GRMZM2G002959 (Tabela 6), cuja anotação cita glutaryl-CoA dehydrogenase como enzima codificada. No entanto, não foi possível identificar sua função em espécies vegetais. O gene GRMZM2G089995 (Tabela 6) foi identificado para a característica 100GW, codificando o fator de transcrição AP2/EREBP transcription factor 209, relacionado às respostas das plantas aos estresses abióticos (Sharoni et al., 2011). Ainda para a característica 100GW, foi identificado o gene GRMZM2G034152, codificando a enzima polyamine oxidase 1, relacionada à resposta ao estresse biótico ou abiótico (Cheng et al., 2017).

Já, para a característica GY, foi identificado um gene (GRMZM2G069618) (Tabela 6), que codifica um domínio de repetição de tetratricopeptídeo (*TPR*) contendo proteína, que está envolvido em respostas da planta a estresse abiótico e biótico e na sinalização hormonal (Sharma e Pandey, 2016), e no desenvolvimento da raiz (Zhang et al., 2015b).

O gene GRMZM2G461936 foi identificado para a característica PE, codificando a proteína argonaute108 (*AGO*). As proteínas *AGO* funcionam cooperando com miRNAs ou siRNAs e regulam o silenciamento gênico no nível pós-transcricional (Zhai et al., 2016), enquanto, para a característica PH, foi identificado o gene GRMZM2G118950, que codifica a proteína ammonium transporter 3, cuja função está relacionada com a absorção de amônio da solução do solo (Sonoda et al., 2003).

Um gene (GRMZM2G086573) (Tabela 6) foi identificado como estando envolvido na característica PV, codificando um fator de transcrição de uma família de proteínas denominada AP2/EREBP transcription factor 24. Os fatores de transcrição AP2/EREBP são encontrados extensivamente em plantas e estão envolvidos no crescimento, desenvolvimento e transdução de sinal em muitas respostas fisiológicas e bioquímicas, como organogênese floral, desenvolvimento de sementes, metabolismo de carbono, resistência a patógenos etc. (Sun et al., 2017).

No ENV2, os genes apresentaram-se distribuídos nos cromossomos 3, 5, 6, 7, 9 e 10 (Tabela 7).

Para a característica EH, foi identificado o gene GRMZM2G002959, codificando a enzima glutaryl-CoA dehydrogenase, cuja função, conforme relatado anteriormente, não foi possível identificar.

O gene GRMZM2G087032 (Tabela 7) foi identificado para a característica 100GW, codificando o putative bifuncional C3H-transcription factor 313, que, de acordo com Liu et al. (2008), pode ser importante durante o estágio inicial do enchimento da semente de milho.

Já, para a característica PE, foi identificado o gene GRMZM2G098793 (Tabela 7), codificando a superfamília de enzimas glycosyltransferase, responsável pelo processo de glicosilação, que é um mecanismo fundamental na determinação da complexidade química e diversidade de produtos naturais vegetais (Wang, 2009), bem como os genes GRMZM2G081048 e GRMZM2G048672 (Tabela 7), o primeiro codificando a enzima oxidoreductase, e o segundo, o fator Macrophage Migration Inhibitory. Oxidoreductases são enzimas que catalisam a transferência de elétrons de uma molécula (reduzidor) para outra (oxidante) e desempenham papéis importantes, não só na transferência de elétrons, mas também em vários processos biossintéticos e vias de biodegradação (Kotera et al., 2009). No entanto, não foi possível identificar a função do fator Macrophage Migration Inhibitory em organismos vegetais.

O gene GRMZM2G043435 foi observado para a característica PH no ENV2, codificando a proteína respiratory burst oxidase-like protein C (RbohC), cuja função foi sugerida por Foreman et al. (2003) como responsável pela expansão das células vegetais, dessa forma, podendo estar envolvida com o desenvolvimento de altura em plantas de milho-pipoca.

Já, para a característica PV, três genes foram identificados, sendo eles: GRMZM2G110726 (Tabela 7), relacionado com a proteína BOBBER 1, encontrado também em *Arabidopsis thaliana* e responsável por funções de desenvolvimento e termotolerância (Perez et al., 2009); e GRMZM2G020150 (Tabela 7), codificando a proteína AP2/EREBP transcription fator 196. Os fatores de transcrição AP2 / EREBP são encontrados extensivamente em plantas e estão envolvidos no crescimento, desenvolvimento e transdução de sinal em muitas respostas fisiológicas e bioquímicas, como organogênese floral, desenvolvimento de sementes, metabolismo de carbono, resistência a patógenos, tolerância a estresse biótico etc. (Sun et al., 2017).

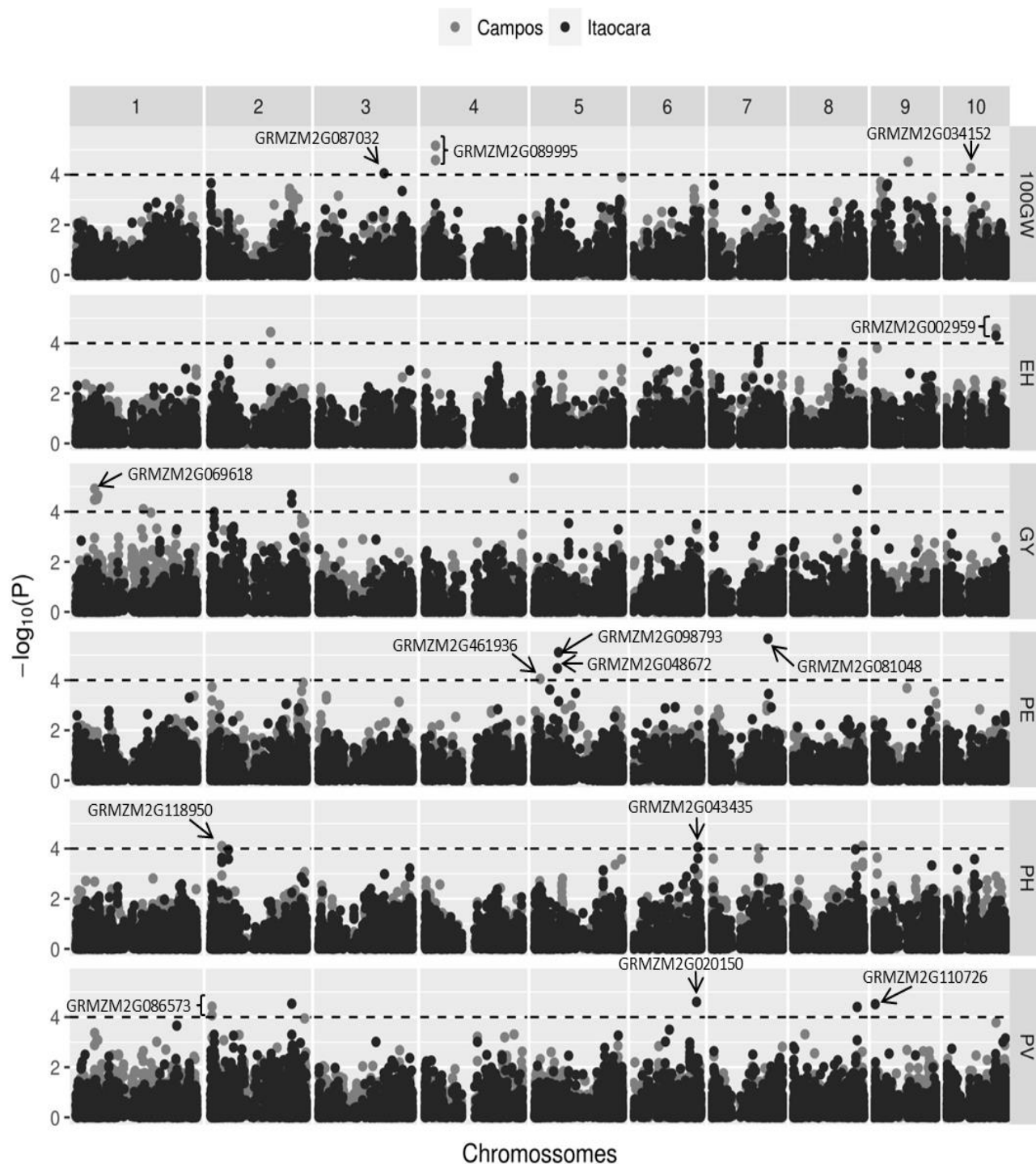


Figura 3. *Manhattan Plot* resultante da análise de associação baseada em modelo linear misto usando marcadores SNPs ($n = 10.507$) nos ambientes de Campos dos Goytacazes (ENV1) e Itaocara (ENV2). Marcadores SNPs significativos são mostrados acima das linhas pontilhadas (valores de $p \leq 0,0001$). As setas indicam os genes candidatos de acordo com o genoma B73, depositado no MaizeGDB.

5.5 Associação via Mapeamento de Herdabilidade Regional

Para os principais achados do RHM em ambos os ambientes, foram identificados possíveis genes candidatos e suas anotações funcionais, conforme relatados nas Tabelas 8 e 9. Nas Figuras 4 e 5, encontram-se os *Manhattan Plots*, evidenciando QTLs significativos ($p \leq 0,001$, ou $-\log_{10}(p) \geq 3$) encontrados nos ENV 1 e ENV 2, respectivamente.

Tabela 8. Genes candidatos obtidos da análise de herdabilidade regional para seis caracteres de interesse em milho-pipoca no ambiente ENV1

Características	Crom.	Menor - Maior Posição (Pb)	Gene ID	Anotação
100GW	4	26378603-26478603	GRMZM5G852329	Serine-threonine protein kinase
			GRMZM2G003642	Mitochondrial pyruvate carrier 2
			GRMZM2G003814	Mitochondrial trans-2-enoyl-coa reductase (MECR, NRBF1)
			GRMZM2G089995	AP2 domain (AP2)
			GRMZM2G090010	Protease family m24 methionyl aminopeptidase, aminopeptidase p
100GW	4	26328603-26428603	GRMZM5G834199	Glucan endo-1,3-beta-glucosidase 5
GY	1	48905387-49005387	GRMZM2G003984	Lon-like ATP-dependent protease
GY	1	56805387-56905387	GRMZM2G069618	TPR domain containing protein
GY	1	56855387-56955387	GRMZM2G353147	GTP diphosphokinase / Stringent fator
GY	6	165623194-165723194	GRMZM5G846343	Protein of unknown function
			GRMZM5G846057	AP2 domain (AP2)
GY	2	225088849-225188849	GRMZM2G037993	Respiratory burst oxidase homolog protein a-related
			GRMZM2G015945	Respiratory burst oxidase homolog protein b

Tabela 8 – Cont.

Características	Crom.	Menor - Maior Posição (Pb)	Gene ID	Anotação			
GY	2	225138849-225238849	GRMZM2G414114	Dnaj domain (dnaj) // TCP family transcription factor (TCP) // Transposase-associated domain			
			GRMZM2G114948	Plant protein of unknown function			
			GRMZM2G023328	Arginine e glutamate-rich protein 1 (ARGLU1)			
			GRMZM2G023585	Hira-interacting protein 3			
GY	4	219528603-219628603	GRMZM2G043242	Zinc finger cw-type coiled-coil domain protein 3			
GY	4	219578603-219678603	GRMZM2G109159	Reticulon-like protein			
PE	2	13388849-13488849	GRMZM5G886913	Predicted membrane protein			
			GRMZM2G051958	Phosphoenolpyruvate carboxykinase ATP			
			GRMZM2G354053	Myosin heavy chain-related //			
			GRMZM5G866405	Isoleucine--tRNA ligase / Isoleucyl-tRNA synthetase			
			GRMZM5G899760	GDP dissociation inhibitor (GDI) // Transcription initiation factor IIA,			
			GRMZM2G059791	2-keto-3-deoxy-l-rhamnonate aldolase			
			GRMZM2G359331	Myosin heavy chain-related			
			AC195235.3_FG003	Phosphoglyceromutase			
			PE	5	13560296-13660296	GRMZM2G461948	Ubiquitin-protein ligase
						AC194618.2_FG008	Dnaj homolog subfamily c member

Tabela 8 – Cont.

Características	Crom.	Menor - Maior Posição (Pb)	Gene ID	Anotação
PE	5	13560296-13660296	GRMZM2G461959	Serine/threonine- protein phosphatase pp2a-1 catalytic
			GRMZM2G461936	Translation initiation factor 2C (eif-2C)
			GRMZM2G161242	Protein Y55F3AM.3, isoform a
			GRMZM2G161222	Serine/threonine protein phosphatase 2a 57 kda regulatory subunit b' alpha isoform
			GRMZM2G148130	Ubiquitin-conjugating enzyme E2 16
PE	5	13610296-13710296	GRMZM2G148098	Homeobox protein transcription factors
			GRMZM2G122185	Pre-mRNA splicing factor
PH	8	171723438-171823438	GRMZM2G133249	Insulysin (IDE, ide)
			GRMZM2G562929	Proteasome subunit alpha type-4
			GRMZM2G133175	Cysteine-rich TM module stress tolerance (CYSTM)
			GRMZM2G133029	Aspartyl protease family protein
			GRMZM2G132991	Conserved oligomeric golgi complex subunit 1
			GRMZM2G132978	Pthr36737:sf1 - expressed protein
				Protein kinase domain (Pkinase) // Salt stress response/antifungal
			GRMZM2G434363	(Stress-antifung)

*Crom: cromossomo

Para a característica 100GW, foram encontrados oito transcritos associados. Na região que corresponde à posição 26378603-26478603 (Mb), verifica-se a presença de cinco possíveis genes candidatos: gene GRMZM5G852329, que codifica a enzima Serine-threonine protein kinase, que demonstrou estar envolvida com a resistência de plantas de arroz ao vírus do enrolamento (RSV) (Lee e Kim, 2015); gene GRMZM2G003642, codificando a proteína Mitochondrial pyruvate carrier 2, cujo gene homólogo NRG1 (gene semelhante a MPC2) da família MPC demonstrou estar envolvido negativamente na sinalização da célula de guarda regulada pelo ácido abscísico (ABA) e na resposta ao estresse hídrico em *Arabidopsis* (Li et al., 2014); gene GRMZM2G003814, responsável pela codificação da enzima Mitochondrial trans-2-enoyl-coa reductase (MECR, NRBF1), que, por sua vez, demonstrou ser responsável por atuar na biossíntese de ácidos graxos em *Euglena gracilis* (Hoffmeister et al., 2005); gene GRMZM2G089995, codificando AP2, envolvido no crescimento, desenvolvimento e transdução de sinal em muitas respostas fisiológicas e bioquímicas, como organogênese floral, desenvolvimento de sementes, metabolismo de carbono, resistência a patógenos etc. (Sun et al., 2017); gene GRMZM2G090010, responsável pela codificação da enzima Protease family m24 methionyl aminopeptidase, aminopeptidase-p, caracterizada pela primeira vez em um organismo vegetal no ano de 2001 (Hauser et al., 2001). No entanto, uma aminopeptidase-p1 (AtAPP1), pertencente à família M24.009 da metalopeptidase M24, foi observada por Murphy et al. (2002) em plântulas de *Arabidopsis thaliana*, e sua expressão foi reconhecida na maioria dos tecidos, incluindo flores, folhas e raízes.

Na região que compreende a posição 26328603–26428603 (pb), ainda no cromossomo 4, foi observado um gene: GRMZM5G834199, que codifica a enzima Glucan endo-1,3-beta-glucosidase 5, relatada por Akiyama et al. (2004) como sendo expressa especificamente em flores e na germinação de sementes de arroz.

Para a característica GY (cromossomo 1, região 48905387-49005387 pb), o gene GRMZM2G003984 foi identificado, o qual codifica a enzima Lon-like ATP-dependent protease, caracterizada por estar envolvida na biogênese mitocondrial, na modulação do metabolismo do carbono, na fosforilação oxidativa e no suprimento de energia, todos pré-requisitos para a germinação de sementes e

estabelecimento de plântulas (Rigas et al., 2014). Ainda nesse cromossomo (posição 56805387–56905387), observou-se a presença do gene GRMZM2G069618, que codifica um domínio de repetição de tetratricopeptídeo (TPR) contendo proteína, envolvido em funções de resposta ao estresse (Sharma e Pandey, 2016) e desenvolvimento (Zhang et al., 2015b). Na posição 56855387 – 56955387 pb (cromossomo 1), observou-se a presença do gene GRMZM2G353147, codificador da enzima GTP diphosphokinase / Stringent factor, necessária para a fertilização de plantas (Masuda et al., 2008).

Ainda para a característica GY, entretanto, no cromossomo 6 (posição 165623194–165723194 pb), foram observados dois genes com possíveis associações à característica: GRMZM5G846343, cuja proteína possui função ainda desconhecida; e GRMZM5G846057, codificando a proteína AP2 domain (AP2), que, como postulado anteriormente, pode estar envolvida em diversas etapas de crescimento e desenvolvimento de tecidos vegetais (Sun et al., 2017).

No cromossomo 2, posição 225088849–225188849 pb, foram identificados dois genes: GRMZM2G037993, responsável pela codificação da enzima Respiratory burst oxidase homolog protein A (rbohA), que demonstrou estar envolvida em diversas funções, como resistência a patógenos (Yoshioka et al., 2003), desenvolvimento de raízes (Arthikala e Quinto, 2018), sendo fundamental para a infecção de *Rhizobium* em *Phaseolus vulgaris* (Arthikala et al., 2017); e GRMZM2G015945, que codifica a enzima Respiratory burst oxidase homolog protein B (rbohB), que, assim como a rbohA, que faz parte do mesma família gênica, também estão envolvidos com a resposta positiva à resistência a patógenos (Yoshioka et al., 2003).

Ainda no cromossomo 2, posição 225138849–225238849 pb, os seguintes genes revelaram significância: i) GRMZM2G414114, proteína Dnaj domain (dnaj) // TCP family transcription factor (TCP) // Transposase-associated domain (Transpos_assoc). Proteínas TCP são fatores de transcrição encontrados exclusivamente em plantas, e estão envolvidos em processos de crescimento e desenvolvimento de uma ampla variedade de plantas (Zheng et al., 2018). Genes TCP foram identificados por Chai et al. (2017), por serem altamente expressos em colmos e espigas de milho, indicando que podem desempenhar funções importantes no crescimento de órgãos axilares e na formação de espigas; ii) GRMZM2G114948, cuja função proteica ainda é desconhecida; iii)

GRMZM2G023328, codificando Arginine and glutamate-rich protein 1 (ARGLU1), que comumente é encontrada em humanos (Zhang et al., 2011; Magomedova et al., 2019), porém não foi possível identificar sua função em organismos vegetais; e iv) GRMZM2G023585, que codifica a proteína Hira-interacting protein 3, cuja função permanece desconhecida.

No cromossomo 4, posição 219528603–219628603 pb, foi possível identificar o gene GRMZM2G043242, codificante da proteína Zinc finger cw-type coiled-coil domain protein (MORC), que demonstrou estar envolvida com funções relacionadas de imunidade de plantas a enfermidades (Manohar et al., 2017). No mesmo cromossomo, mas na posição 219578603–219678603 pb, o gene GRMZM2G109159 foi identificado, o qual codifica a proteína Reticulon-like protein, localizada no retículo endoplasmático de plantas (Tolley et al., 2008) e que possui importante função no tráfego celular, cujo processo metabólico tem relevante envolvimento no preenchimento de sementes (Nziengui e Schoefs, 2009).

Para a característica PE, foram identificados 24 genes em diferentes cromossomos. No cromossomo 2 (posição 13388849–13488849 pb), os seguintes genes foram observados: i) gene GRMZM5G886913, que codifica a proteína Predicted membrane protein. As proteínas de membrana são importantes em muitas funções, devido às suas localizações celulares (Tai et al., 2011); ii) gene GRMZM2G051958, responsável pela codificação da enzima Phosphoenolpyruvate carboxykinase ATP (PECC). Um gene *pepc*, clonado a partir de milho, foi capaz de conferir tolerância à seca e aumentar o rendimento de grãos em trigo transgênico, bem como aumentar a taxa fotossintética de plantas dessa mesma cultura, aumentando o peso da semente por espiga e o peso de mil grãos (Hu et al., 2012; Qin et al., 2016); iii) gene GRMZM2G354053, codificando Myosin heavy chain-related. Alguns membros da família das miosinas estão envolvidos, direta ou indiretamente, no movimento do complexo de Golgi e das mitocôndrias nas células vegetais (Avisar et al., 2009); iv) gene GRMZM5G866405, codificante da enzima Isoleucine--tRNA ligase / Isoleucyl-tRNA synthetase, que é uma aminoacil-tRNA sintetase (Kermgard et al., 2017); no entanto, não foi possível identificar sua função em plantas; v) gene GRMZM5G899760, responsável pela codificação da proteína GDP dissociation inhibitor (GDI), associada à germinação do pólen e ao crescimento do tubo

polínico em *Arabidopsis thaliana* (L.) Heynh. (Ge et al., 2011); vi) gene GRMZM2G059791, codificando a enzima 2-keto-3-deoxy-l-ramnonate aldolase, que já foi observada em vários organismos vegetais, como *Fragaria vesca* subsp. *Vesca* (Zhang et al., 2016b) e *Vitis vinífera* L. (Li et al., 2015); no entanto, não foi possível identificar a função que desempenha nesses; vii) gene GRMZM2G359331, codificando a proteína Myosin heavy chain-related, que, conforme relatado anteriormente, pode estar envolvida no movimento do complexo de Golgi e das mitocôndrias nas células vegetais (Avisar et al., 2009); e viii) gene AC195235.3_FG003, responsável pela codificação da enzima Phosphoglyceromutase, que demonstrou atuar no aumento de concentração em raízes de *Cucumis sativus* L. deficientes a ferro (Donnini et al., 2010).

No cromossomo 5, posição 13560296 – 13660296, foram encontrados os genes: i) GRMZM2G461948, codificando a enzima Ubiquitin-protein ligase E3 B (UBE3B), cuja função, em organismos vegetais, não foi possível identificar; ii) AC194618.2_FG008, que codifica Dnaj homolog subfamily c member, relacionada com a tolerância ao frio em plantas de tomate transgênicos (Kong et al., 2014); iii) GRMZM2G461959, que codifica Serine/threonine-protein phosphatase pp2a-1 catalytic, altamente expressa em caules, flores e raízes de *Oryza sativa* (Yu et al., 2003); iv) GRMZM2G461936, responsável pela codificação do Translation initiation factor 2C (eif-2C). Os fatores de iniciação da tradução (eIFs), além da síntese de proteínas, regulam vários aspectos do desenvolvimento da planta e sua interação com o meio ambiente (Dutt et al., 2015); v) GRMZM2G161242, codificando Protein Y55F3AM, isoform a, cuja função não foi possível identificar; vi) GRMZM2G161222, que codifica Serine/threonine protein phosphatase 2a 57 kDa regulatory subunit b' alpha isoform. As enzimas Serine/threonine protein phosphatase 2A (PP2As) participam no controle de respostas de estresse biótico e abiótico em plantas (País et al., 2009); e vi) GRMZM2G148130, codificando Ubiquitin-conjugating enzyme E2 16. Níveis de Ubiquitin-conjugating enzyme E2s podem aumentar ou diminuir, quando plantas de milho são submetidas a diferentes tipos de estresse abiótico, indicando as funções críticas dessa família de genes na manutenção do crescimento normal do milho sob condições de estresse (Jue et al., 2015).

Ainda no cromossomo 5, posição 13610296 – 13710296, foram observados dois genes: GRMZM2G148098, codificando Homeobox protein transcription

factors, que desempenha um papel importante no estabelecimento e na manutenção do equilíbrio hormonal em plantas de milho, fundamental para a manutenção apropriada do meristema e da iniciação de órgãos (Bolduc et al., 2012); e GRMZM2G122185, codificando Pre-mRNA splicing factor, envolvido na sinalização de temperatura em plantas (Capovilla et al., 2015).

Para a característica PH, no cromossomo 8, posição 171723438 – 171823438, foram identificados os seguintes genes: i) GRMZM2G133249, codificando Insulysin/IDE, que é uma enzima degradadora de insulina e está envolvida na degradação de beta-amiloide (A β). Acredita-se que o acúmulo cerebral de proteína A β desempenha papel importante na patogênese da doença de Alzheimer (Kim et al., 2007; Zuo e Jia, 2009); ii) GRMZM2G562929, que codifica Proteasome subunit alpha type-4, cuja função não foi possível identificar; iii) GRMZM2G133175, codificando Cysteine-rich TM module stress tolerance (CYSTM), que foi responsável por conferir tolerância a metais pesados, como cádmio e cobre em *Digitaria ciliaris* e *Oryza sativa* (Kuramata et al., 2009) e em espécies de eucariotos (Venancio e Aravind, 2010); iv) GRMZM2G133029, responsável pela codificação da proteína Aspartyl protease family protein, que, associada a uma cochaperona molecular, desencadeia autofagia e defesa da planta a fungos (Li et al., 2016); v) GRMZM2G132991, que codifica Conserved oligomeric golgi complex subunit 1, o qual demonstrou estar envolvido com a resistência à penetração da cevada pelo fungo do míldio da cevada (Ostertag et al., 2013); vi) GRMZM2G132978, que codifica uma proteína expressa, no entanto, sua identificação não foi possível; e vii) GRMZM2G434363, que codifica uma Protein kinase domain (Pkinase) // Salt stress response/antifungal (Stress-antifung), a qual pode ter função importante na resposta ao estresse salino em plantas de arroz (Zhang et al., 2009).

Tabela 9. Genes candidatos obtidos da análise de herdabilidade regional para seis caracteres de interesse em milho-pipoca no ambiente ENV2

Características	Crom.	Menor - Maior Posição (Pb)	Gene ID	Anotação
100GW	6	155323194-155423194	GRMZM2G010357	CCR4-NOT transcription complex subunit 2 (CNOT2, NOT2)
			GRMZM2G462717	60S Ribosomal protein l34
			GRMZM2G701218	Myb_DNA-binding) // Myb_CC_LHEQLE)
			GRMZM2G157246	Ring zinc finger protein
			GRMZM2G157263	Cytochrome-b5 reductase
			GRMZM2G157267	Inositol 5-phosphatase
EH	7	115031917-115131917	GRMZM2G022095	rRNA N-glycosylase / rRNA N-glycosidase
EH	7	115081917-115181917	GRMZM5G837058	Golgi SNAP receptor complex member 1-1
			GRMZM2G071059	CCR4-NOT transcription complex subunit 7/8 (CNOT7_8, CAF1, POP2)
EH	4	178478603-178578603	GRMZM2G170313	Prolyl-tRNA synthetase associated domain-containing protein 1-related
			GRMZM2G473016	Ring zinc finger protein
			GRMZM2G060630	Solute carrier family 25 (mitochondrial phosphate transporter), member 3
			GRMZM2G060554	Remorin, C-terminal region (Remorin_C)
			GRMZM2G356046	Mannan endo-1,4-beta-mannosidase (MAN)
GY	2	200888849-200988849	GRMZM2G024622	RNA polymerase II ctd phosphatase
PV	9	3856799-3956799	GRMZM2G048243	Serine-threonine protein kinase
			GRMZM2G110726	BOBBER 1
PV	2	200888849-200988849	GRMZM2G024622	RNA polymerase ii ctd phosphatase

*Crom: cromossomo

Para a característica 100GW, foram observados genes apenas no cromossomo 6 (posição 155323194-155423194 pb). São os seguintes: i) GRMZM2G010357, que codifica CCR4-NOT *transcription complex subunit 2* (CNOT2, NOT2), os quais são reguladores transcricionais gerais, e demonstraram ser essenciais para o desenvolvimento das plantas de arroz (*Oryza sativa*) e *Arabidopsis* (Wang et al., 2013); ii) GRMZM2G462717, responsável por 60S RIBOSOMAL PROTEIN L34, que pode desempenhar uma função negativa nos

processos metabólicos de adaptação de sementes de soja transgênicas ao estresse de baixa temperatura durante a embebição a baixa temperatura (Cheng et al., 2010); iii) GRMZM2G701218, codificando MYB_DNA-binding // MYB_CC_LHEQLE. Os fatores de transcrição MYB são um grupo de fatores de transcrição, que contém um domínio conservado chamado MYB_DNA-binding. Um gene OsMPH1, que codifica um fator de transcrição MYB, relacionado à regulação da altura de plantas de arroz, demonstrou também estar envolvido na melhoria de rendimento de grãos dessa mesma espécie (Zhang et al., 2017); iv) GRMZM2G157246, responsável pela codificação da proteína RING zinc finger protein. Um gene de arroz, que codifica uma proteína RING zinc-finger, demonstrou ser responsável pelo aumento da abertura dos estômatos da espécie (Hsu et al., 2014); v) GRMZM2G157263, que codifica cytochrome-b5 reductase, necessária para o funcionamento correto do pólen e para a maturação das sementes de *Arabidopsis thaliana* (Wayne et al., 2013); vi) GRMZM2G157267, codificando inositol 5-phosphatase. Plantas de *Arabidopsis thaliana* transgênica expressando o Tipo 1 de Inositol 5-Fosfatase demonstram aumento da tolerância à seca (Perera et al., 2008).

Para a característica EH, posição (115031917–115131917 pb), foi identificado apenas o gene GRMZM2G022095, que codifica a enzima rRNA N-glycosylase/rRNA N-glycosidase, também chamada de enzima inativadora de ribossomos (RIPs) (Sharma et al., 2004), possuindo atividade antiviral, antifúngica e inseticida, e sua expressão nas plantas é aumentada sob condições estressantes (Stirpe, 2013).

Para a mesma característica e mesmo cromossomo, entretanto na posição 115081917–115181917 pb, foram identificados dois genes, sendo esses: i) GRMZM5G837058, que codifica Golgi SNAP receptor complex member 1-1, envolvida no transporte entre complexo de Golgi e retículo endoplasmático (Tian et al., 2018); e ii) GRMZM2G071059, codificando CCR4-NOT transcription complex subunit 7/8 (CNOT7_8, CAF1, POP2). *AtCAF1a* e *AtCAF1b*, homólogos putativos de *Arabidopsis* do gene *CAF1* de levedura, demonstraram estar envolvidos na mediação de respostas ao estresse abiótico em *Arabidopsis*, porém não são responsivos a todos os estresses ambientais (Walley et al., 2010).

No cromossomo 4 (posição 178478603–178578603 pb), foram identificados: i) gene GRMZM2G170313, que codifica Prolyl-tRNA synthetase

associated domain-containing protein 1-related, cuja função, em organismos vegetais, não foi possível identificar; ii) gene GRMZM2G473016, responsável pela codificação da proteína Ring zinc finger, que, conforme relatado anteriormente, está envolvido com o aumento da abertura estomática em plantas de arroz (Hsu et al., 2014); iii) gene GRMZM2G060630, que codifica a proteína mitochondrial phosphate transporter member 3, que demonstrou ser indispensável para o crescimento e desenvolvimento normal da *Arabidopsis* (Jia et al., 2015); iv) gene GRMZM2G060554, codificando Remorin, C-terminal region (Remorin_C). Um gene Remorin (*ZmREM6.3*) presente na cultura do milho (*Zea mays* L.) foi associado à resistência de genótipos dessa cultura à helmintosporiose, causada pelo fungo *Setosphaeria túrcica* (Jamann et al., 2016); e v) gene GRMZM2G356046, responsável pela síntese da enzima Mannan endo-1,4-beta-mannosidase (MAN). Iglesias-Fernández et al. (2011) sugeriram que β -mannosidase é importante para a germinação de sementes de *Arabidopsis thaliana*.

Para a característica GY, apenas um gene foi identificado, presente na região entre 200.888.849 – 200.988.849 pb. Trata-se de GRMZM2G024622, responsável pela codificação de uma RNA polymerase II ctd phosphatase, que demonstrou estar envolvida no aumento da tolerância ao estresse térmico em plantas de Crisântemo (*Chrysanthemum morifolium*) (Qi et al., 2018).

Com relação à característica PV, foram observados genes em distintas posições e cromossomos. No cromossomo 9, posição 3856799 – 3956799 pb, foram identificados os seguintes genes: i) GRMZM2G048243, codificando uma Serine-threonine protein kinase, que, conforme relatado anteriormente, está relacionada com a resistência de plantas de arroz ao RSV (Lee e Kim, 2015); e ii) GRMZM2G110726, codificando BOBBER1, que está envolvida em funções de desenvolvimento e termotolerância (Perez et al., 2009).

Na região 200888849 – 200988849 do cromossomo 2, foi observada a presença do gene GRMZM2G024622, que, conforme já aqui relatado, codifica uma RNA polymerase II CTD phosphatase, envolvida na resposta ao estresse térmico em plantas de crisântemo (Qi et al., 2018).

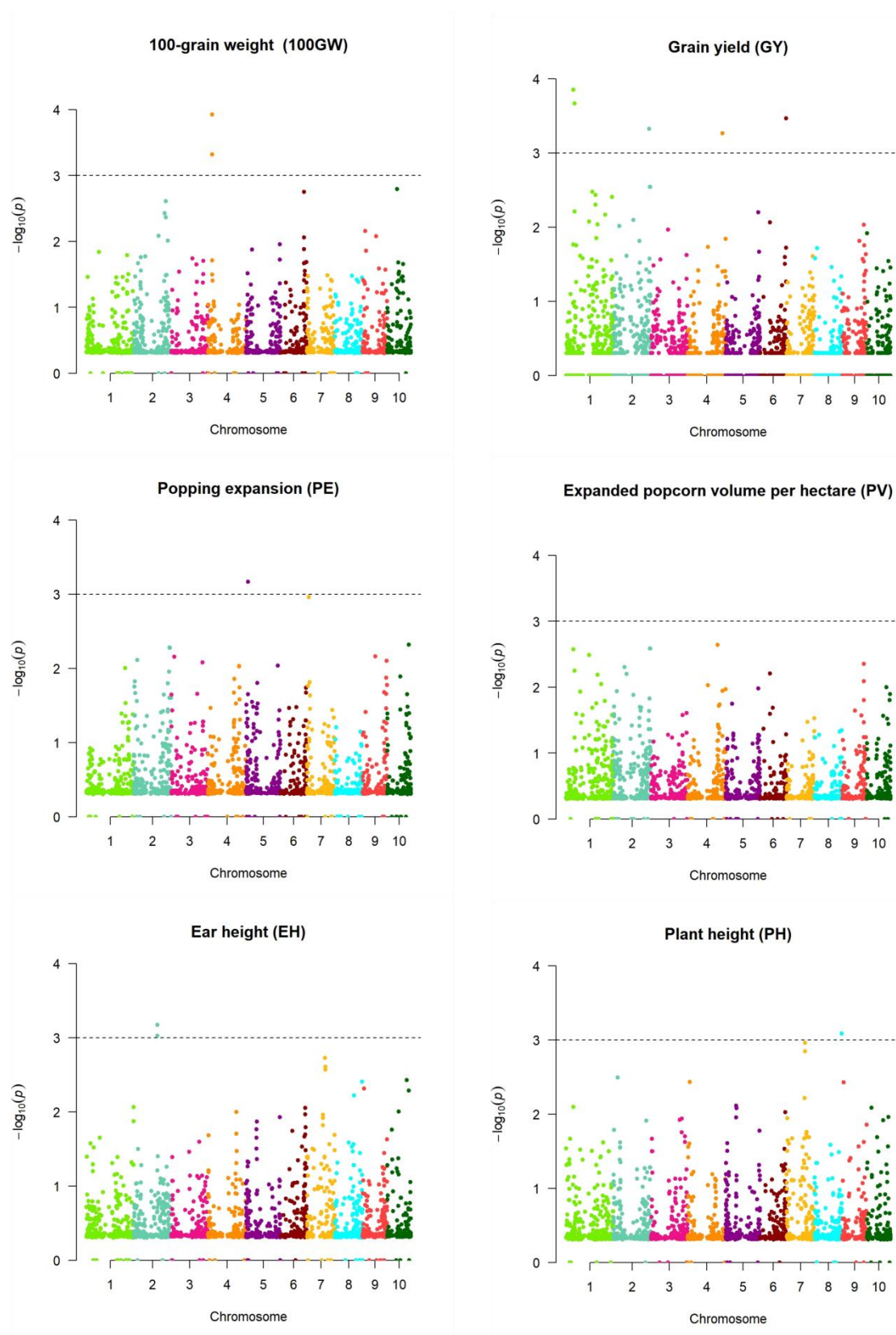


Figura 4. *Manhattan Plot* resultante da análise RHM utilizando marcadores SNP ($n = 10.507$) no ambiente de Campos dos Goytacazes (ENV1). Regiões significativas são mostradas acima das linhas pontilhadas (valor de $p \leq 0,001$).

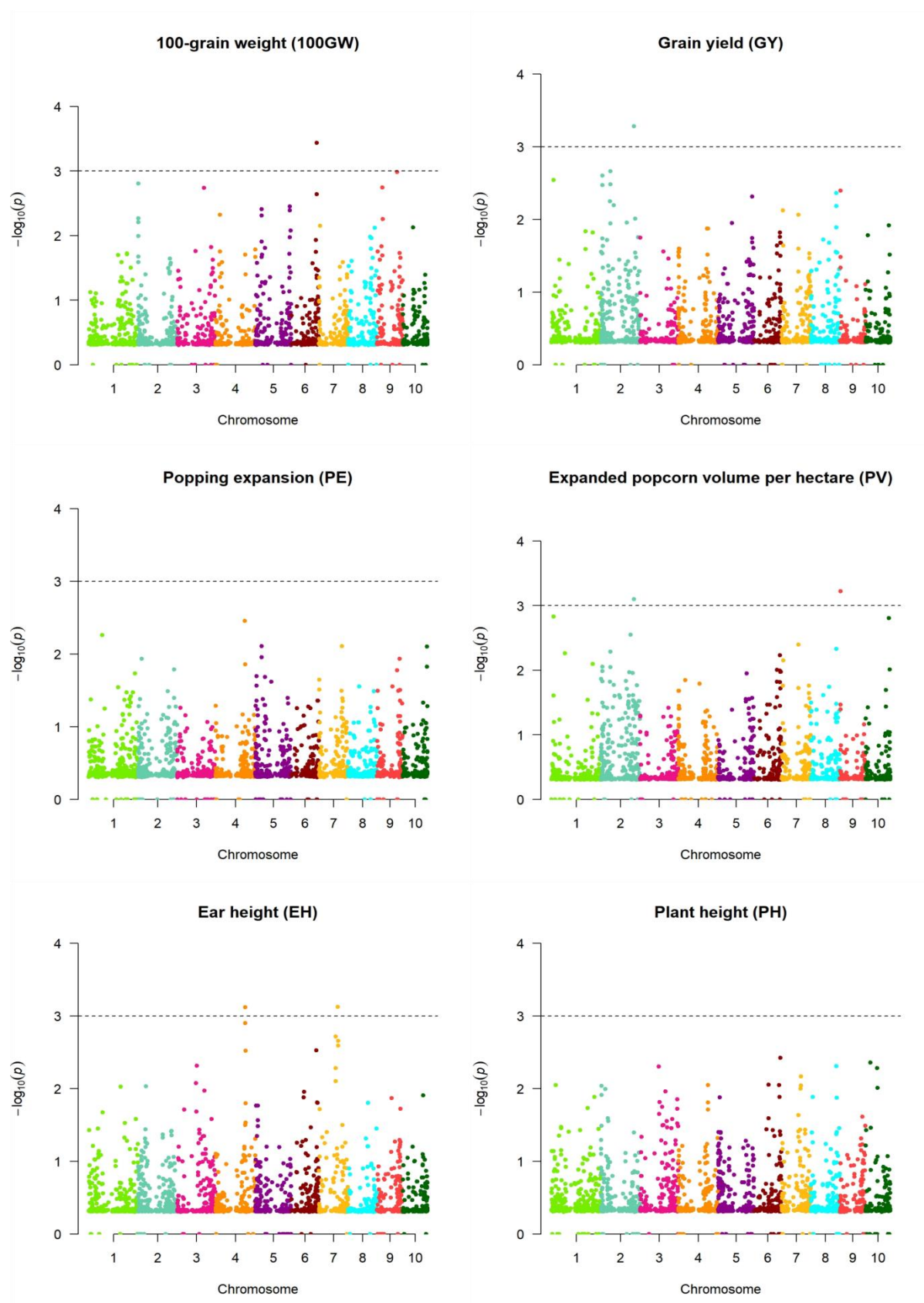


Figura 5. *Manhattan Plot* resultante da análise RHM utilizando marcadores SNP ($n = 10.507$) no ambiente de Itaocara (ENV2). Regiões significativas são mostradas acima das linhas pontilhadas (valor de $p \leq 0,001$).

5.6 Análise comparativa dos resultados da Associação de Modelo Linear Misto (MLMA) e Mapeamento de Herdabilidade Regional (RHM)

Uma maior proporção de associações genômicas foi observada na abordagem RHM (43 no ENV1 e 19 no ENV2), quando comparada com a abordagem MLMA (7 no ENV1 e 8 no ENV2).

Das sete associações detectadas pela MLMA (Tabela 6) no ENV1, três também foram detectadas na análise RHM (Tabela 8) nos mesmos intervalos. Foram elas: gene GRMZM2G089995, relacionado com a característica 100GW; gene GRMZM2G069618 associado com GY; e gene GRMZM2G461936, relacionado com a característica PE. No entanto, no ENV2, das oito associações observadas na MLMA (Tabela 7), apenas uma foi detectada na análise RHM (Tabela 9), sendo esta associada à característica PV (gene GRMZM2G110726).

No geral, ambas as análises tendem a identificar as mesmas regiões genômicas, embora o RHM possibilite uma análise mais detalhada, e, por isso, identifica mais genes nas regiões significativas. Observa-se, ainda, que houve um número maior de SNPs associados às características PE, GY e PH, no ENV1, e à característica EH no ENV2. Em ambos os ambientes, o número de SNPs variou em todas as características, não sendo possível afirmar qual delas possui uma arquitetura genética mais complexa.

6. DISCUSSÃO

Este estudo teve como objetivo identificar regiões genômicas subjacentes a características relacionadas à produtividade e ao crescimento em genótipos de milho-pipoca e comparar os valores de herdabilidades observados nos métodos MLM e RHM. Ambas as análises demonstraram identificar regiões importantes para a elucidação a respeito dos genes envolvidos na determinação das características complexas da espécie; no entanto, o RHM permitiu a identificação de maior número de regiões, resultando em maior quantitativo de genes candidatos identificados. Sem embargo, esperavam-se valores de herdabilidade mais elevados pelo método RHM, quando comparado com a análise MLMA. Atualmente, esse é o primeiro estudo no qual o RHM é aplicada em milho-pipoca.

Estudos de associação genômica na espécie foram desenvolvidos por Paes et al. (2016), que identificaram QTLs envolvidos na qualidade da pipoca, em populações de germoplasma tropical e temperado, oriundas de distintos ciclos de seleção; Senhorinho et al. (2019), identificaram quatro SNPs significativamente associados a capacidade de expansão, em linhagens de milho-pipoca, de forma que os genes anotados, subjacentes aos SNPs, mostraram-se relacionados com funções associadas ao conteúdo de amido; e Coan et al. (2018) que, em estudo envolvendo linhagens de milho-pipoca e milho tropical, identificaram 14 SNPs significativamente associados à podridão de espiga, causada por *Fusarium verticillioides* (Sacc.) Nirenberg.

Esses estudos foram realizados utilizando linhagens, enquanto nosso objetivo foi identificar regiões de interesse em genótipos oriundos de ciclos de seleção recorrente intrapopulacional, que demonstram apresentar maior grau de dificuldade de identificação, uma vez que Guimarães et al. (2018), avaliando os diversos ciclos de seleção recorrente (C0 a C8) da mesma população de milho-pipoca, verificaram uma redução no número de heterozigotos à medida que os ciclos avançavam, implicando, conseqüentemente, a redução da diversidade genética da população. Embora haja menor variabilidade genética nesse tipo de material, as associações encontradas demonstram ser mais úteis para a utilização na prática pelos melhoristas (Resende et al., 2017).

Resende et al. (2017) realizaram estudos de genome-wide association studies (GWAS) e RHM para sete características de crescimento, resistência à madeira e doenças em uma população de híbridos de *Eucalyptus* e verificaram que o GWAS detectou 13 SNPs com associações significativas em todo o genoma, enquanto o RHM detectou 26 locos de características quantitativas (QTLs), abrangendo 2.191 polimorfismos de nucleotídeo único (SNPs), capturando frações maiores de herdabilidade de características quando comparado o GWAS. Ainda segundo os autores, alelos de baixa frequência em populações naturais tornam-se consideravelmente mais comuns quando amostrados em populações melhoradas, e, considerando que são identificados por abordagens de RHM, podem ser facilmente rastreados em gerações de melhoramento.

6.1 Caracterização da população de estudo

A metade do decaimento de LD variou entre os cromossomos (76-151Kb). O desequilíbrio de ligação em espécies alógamas, como o milho, decai em distâncias relativamente mais curtas do que em espécies autógamas (Newell et al., 2011). No entanto, Morosini et al. (2017), avaliando sessenta e quatro linhagens de milho, calcularam o LD do genoma e observaram que o comprimento médio do decaimento LD em todos eles foi de 80-100 kb. A média de desequilíbrio de ligação para todos os cromossomos foi estimada em ~ 200 kb em 144 linhagens milho submetidas à análise de GWAS para identificação de genes relacionados à resistência a *Sphacelotheca reiliana* (Kühn) Clint (Wang et al., 2012).

Os resultados pelo critério de Evanno et al. (2005) indicaram que a população apresenta estratificação genética. Esses resultados podem ser explicados pelo parentesco genético dos indivíduos, uma vez que os resultados do modelo de inferência sobre subpopulações de Pritchard et al. (2000) estão em consonância com a dispersão dos coeficientes de parentesco reportada pelos componentes principais.

Em estudo prévio com essa população, Schwantes et al. (2020, no *prelo*) mostram que a maior parte do parentesco calculado pelos SNPs é de baixa magnitude. Portanto, esses resultados apontam que a população de estudo não apresenta elevada estratificação e que essa possível estrutura observada deve ser explicada por grupos de indivíduos que aleatoriamente receberam alguns alelos em comum na recombinação; dessa forma, o $K = 2$ apresenta ser uma boa alternativa para a correção dessa estrutura existente.

O controle da estrutura da população é um procedimento padrão, embora, quando as características sejam fortemente confundidas, possa haver uma redução no poder da análise. Esse problema pode surgir ao estudar características, como tempo de floração e tolerância ao frio, que são influenciadas por gradientes ambientais, os quais se sobrepõem aos padrões de estrutura da população (Brachi et al., 2011).

Métodos baseados no MLM têm sido utilizados para corrigir a relação genética e a estrutura populacional, uma vez que eles levam em consideração a estrutura da população e a relação familiar (Zeng et al., 2017a). A estrutura da população foi corrigida com sucesso em diversos estudos com milho (Weng et al., 2011; Wang et al., 2012; Wen et al., 2014; Liu et al., 2016), dessa forma, a correção da estrutura populacional demonstrou ser um procedimento essencial para estudos de associação.

6.3 Comparação de herdabilidades genômicas da MLMA e do RHM

As estimativas de herdabilidade por RHM foram inferiores para a maioria das características (Tabelas 4 e 5), quando comparadas com os valores de herdabilidade obtidos pela análise MLMA (Tabela 2).

Por RHM no ENV1, para GY — uma das principais características de interesse na cultura do milho-pipoca —, uma série de valores de herdabilidade foram obtidos, o que pode ser devido a efeitos populacionais, fatores ambientais

ou precisão experimental (Resende et al., 2018), ou, mais provavelmente, devido ao particionamento do genoma que ocorre na metodologia RHM. Os valores foram estimados em uma faixa de 10 a 90%, observando-se, ainda, que estes foram maiores no RHM (Tabela 4) em comparação ao MLMA, cujo valor observado foi inferior a 1% (Tabela 2). Valores elevados de herdabilidade para peso de grãos por planta, em genótipos de milho, foram observados por Li et al. (2011), que constataram uma herdabilidade de 89% para a característica. Além disso, Viana et al. (2017), simulando dados de linhagens parentais, genótipos e fenótipos F_1 e F_2 e utilizando a abordagem de máxima verossimilhança por mapeamento de intervalos (Lander e Botstein, 1989) para QTLs de baixa herdabilidade e alta densidade de SNPs, obtiveram um valor de aproximadamente 95% de herdabilidade para a característica rendimento de grãos, quando da utilização de um número amostral igual a 200.

Estudos relatam que rendimento de grãos é uma característica que apresenta herdabilidade variando de baixa a moderada (Arnhold et al., 2009; Soni e Khanorkar, 2014), indicando um menor controle genético aditivo e elevada influência ambiental. No entanto, Rodovalho et al. (2008) e Ribeiro et al. (2012) relatam que a herdabilidade dessa característica demonstra ser alta e significativa, sugerindo, assim, que esses valores dependem da população submetida à análise. Ribeiro et al. (2016) observaram que a herdabilidade se manteve estável nos ciclos 4, 5 e 6 (45,97%, 51,94% e 45,04%, respectivamente) no programa de seleção recorrente intrapopulacional da população UENF 14. Além disso, Guimarães et al. (2019), ao calcularem a herdabilidade de GY no ciclo 9 da mesma população, constataram valor de 91%.

Resende et al. (2018) observaram que a herdabilidade genômica estimada capturou uma proporção relativamente grande (72%) da herdabilidade total das características, sendo de grande valor na identificação de alelos favoráveis ao rendimento de grãos de *Phaseolus vulgaris* L. Este fato sugere que a análise via MLMA não foi capaz de capturar a variância genética da característica, enquanto o RHM apresentou eficácia na detecção da porção herdável da característica na população.

Nota-se que as características que apresentaram maiores valores de herdabilidade na abordagem MLMA no ENV1 (Tabela 2) — EH e PH — foram as que apresentaram os menores valores de herdabilidade no RHM (Tabela 4).

Altura de espiga (EH) e altura de planta (PH) apresentaram valores de herdabilidade de 70% e 80%, respectivamente, quando foram avaliadas famílias $F_{2:3}$, resultantes de cruzamento entre linhagens de milho-pipoca e linhagens de milho dentado (Li et al., 2007).

Isso sugere que os valores observados neste estudo ocorreram devido à falta de poder do método RHM em capturar boa parte da variância genética da característica. Caballero et al. (2015) afirmam que caracteres de baixa herdabilidade não apresentam valores elevados de herdabilidade regional. No entanto, observou-se que EH e PH (ENV1) apresentaram valores moderados de herdabilidade na abordagem MLMA, levando a sugerir que ambas as características possuem alta herdabilidade, quando da avaliação do genoma completo. Entretanto, ao fracionar o genoma para a realização do RHM, o método não é capaz de capturar a herdabilidade presente nas janelas genômicas, provavelmente, por serem de pequeno efeito, o que resulta na não contabilização pela análise RHM. Resende et al. (2018) sugerem que a proporção da herdabilidade não amostrada é devido ao LD entre marcadores e genes e/ou variantes raras de alelos não amostrados.

Não obstante, no ENV2, os valores de herdabilidade foram considerados baixos em ambas as análises. Com exceção da característica PE, na abordagem MLMA (Tabela 2), todas as outras demonstraram apresentar baixa proporção herdável da variabilidade total (Tabela 5). Estudos anteriores relatam que a capacidade de expansão dos grãos de milho-pipoca possui ação gênica aditiva (Larish e Brewbaker, 1999; Pereira e Amaral Júnior, 2001; Ribeiro et al., 2016; Schwantes et al., 2018).

Sugere-se, assim, que ambos os métodos são capazes de detectar a porção herdável das principais características da cultura. No entanto, o método de análise via modelo linear misto (MLMA) comprovou ser mais eficaz para determinadas características, tais como EH, PH e PE, enquanto o mapeamento de herdabilidade regional (RHM) demonstrou ser mais eficaz para a característica GY.

Algumas hipóteses podem ser sugeridas quanto a isso, como o fato das características PE e PH estarem diretamente correlacionadas e possuírem alta herdabilidade (Li et al., 2007), porém, ao particionar o genoma, como ocorre no método RHM, os valores dentro de cada janela genômica podem não ser grandes

o suficiente a ponto de serem capturados pela análise, enquanto, na abordagem MLMA, há um somatório desses valores, fazendo com que a herdabilidade seja apresentada de forma mais abundante, assim ocorrendo também com PE. Fato esse que não ocorre com a característica GY, cujos maiores valores de herdabilidade foram observados no RHM, que, apesar de demonstrar ser uma característica altamente influenciável pelo ambiente (Arnhold et al., 2009; Soni e Khanorkar, 2014), a população aqui amostrada demonstrou manter-se estável com o avanço dos ciclos de seleção recorrente intrapopulacional (Ribeiro et al., 2016; Guimarães et al., 2019).

6.3 Associação de Modelo Linear Misto (MLMA)

Vários genes podem estar presentes no intervalo de uma região associada, isto é, perto de SNPs, e estes podem ser identificados por MLMA. Esta análise identificou sete e oito genes relacionados às características estudadas no ENV1 e ENV2, respectivamente. O gene GRMZM2G002959, identificado para EH nos dois ambientes, codifica a enzima glutaril-CoA desidrogenase, mas sua função em espécies de plantas não foi identificada.

O gene GRMZM2G089995, identificado para 100GW no ENV1, codifica o transcription factor 209, um membro da família de proteínas AP2/EREBP, que está diretamente envolvido na resposta a estresses bióticos e abióticos (Dietz et al., 2010; Mun et al., 2017). Para a mesma característica, o gene GRMZM2G034152 codifica PAO1, cuja função também está relacionada a respostas a estresses abióticos em arroz (Cheng et al., 2017) e a estresses bióticos em algodão (Mo et al., 2015). Isso indica que a característica 100GW está relacionada a respostas de estresses (biótico e abiótico). Este resultado corrobora os resultados anteriores obtidos para arroz submetido a estresse hídrico, no qual foi observada uma redução no peso de 1000 grãos (Zain et al., 2014), bem como a premissa de que o estresse hídrico contribuiu para redução na produtividade em linhagens de arroz (Terra et al., 2013). Em soja, sementes cultivadas sob condições climáticas adversas também apresentaram menor massa para 100 grãos (Borrmann et al., 2009).

No ENV2, o gene candidato identificado para a característica 100GW (GRMZM2G087032) corresponde à proteína putative bifunctional C3H-transcription factor 313, que pode ser importante durante o estágio inicial do

enchimento da semente de milho (Liu et al., 2008). No entanto, o estudo dessa proteína ainda está em um estágio inicial, já que a pesquisa sobre esse assunto é muito limitada.

A característica GY (ENV1) demonstrou relação com o gene GRMZM2G069618, que, por sua vez, codifica TPR domain containing protein, a qual desempenha várias funções nos organismos vegetais, que vão desde o desenvolvimento (Awasthi et al., 2012; Zhang et al., 2015b) à resposta ao estresse hormonal (Sharma e Pandey, 2016).

No ENV1, o gene GRMZM2G461936, que codifica a proteína argonaute108, foi identificado como estando relacionado à PE (capacidade de expansão). As proteínas Argonaute (AGO) são fundamentais na regulação da expressão gênica e são essenciais para vários processos de desenvolvimento (Zhai et al., 2014). Esses autores identificaram o gene ZmAGO18b em genótipos de milho e observaram alta expressão dela em tecidos reprodutivos. Embora nenhum estudo tenha identificado especificamente a proteína argonaute108 em plantas, observou-se que a deficiência de uma proteína Argonaute, denominada AGO10, induziu desenvolvimento anormal do meristema apical de gemas em plantas de *Arabidopsis* (Lynn et al., 1999).

Também para a PE, o gene GRMZM2G098793 foi identificado como codificador da superfamília da enzima glicosiltransferase no ENV2. Em amido de endosperma de trigo, observou-se que o arabinosilano, responsável por uma grande quantidade de polissacarídeo, foi reduzido ao suprimir dois genes homólogos de glicosiltransferase (Lovegrove et al., 2013). Isso permitiu concluir que o gene identificado neste estudo também está envolvido com a síntese de amido em grãos de milho, e que sua supressão pode modificar o endosperma, cuja função está diretamente relacionada com a capacidade de expansão do grão sob altas temperaturas.

O gene GRMZM2G081048, também associado à PE, codifica a enzima oxidoreductase. Neste estudo, não foi possível identificar qual tipo de oxidoreductase está envolvido com a característica, prejudicando uma análise mais aprofundada de suas funções. O gene GRMZM2G048672, também identificado como relacionado à PE, mostrou ser responsável pela citocina imunorreguladora, denominada fator inibitório da migração de macrófagos. Entretanto, nenhum relato de sua função em plantas foi encontrado na literatura.

Em ENV1, o gene GRMZM2G118950, identificado pela característica PH, codifica o ammonium transporter 3 (amt3). Ammonium transporters (amt) são responsáveis pela absorção de amônio da solução do solo (Howitt e Udvardi, 2000; Sonoda et al., 2003). Dois Amts localizados na rizoderme (ZmAMT1; 1a e ZmAMT1; 3) foram identificados, concluindo que estes são, provavelmente, os principais componentes do sistema de transporte de amônio de alta afinidade em raízes de milho (Gu et al., 2013). No milho, a nutrição com amônio confere vários efeitos benéficos, como o aumento da densidade e extensão da raiz (Bloom et al., 2002). Portanto, esse fator pode influenciar diretamente a PH das plantas de milho-pipoca da população em estudo.

No ENV2, o gene GRMZM2G043435 foi associado à PH, codificando a proteína respiratory burst oxidase-like protein C (RbohC). Em *Arabidopsis*, os mutantes deficientes em RbohC tinham pelos radiculares curtos em raízes atrofiadas, sugerindo que essa espécie de Rboh regula a expansão de células vegetais (Foreman et al., 2003) e pode estar envolvida no desenvolvimento da altura de plantas de milho-pipoca.

Um gene — GRMZM2G086573 — foi identificado como relacionado à expressão da característica PV no ENV1, codificando um fator de transcrição de uma família de proteínas chamada AP2-EREBP-transcription factor 24. A característica distintiva das proteínas AP2/EREBP é que esta família contém um ou dois domínios APETALA2 (AP2). Observou-se que o gene regulador do amido de arroz (RSR1), um fator de transcrição da família APETALA2, regula negativamente a expressão de genes de síntese de amido tipo I, e que a deficiência de RSR1 resulta em uma expressão aumentada de genes de síntese de amido em sementes da espécie (Fu e Xue, 2010). O volume de pipoca é dependente da umidade do grão e do teor de amido, que é convertido em vapor e exerce pressão sobre o endosperma (Singh et al., 2017).

Em ENV2, dois genes foram identificados: GRMZM2G110726, que codifica a proteína BOBBER1 (domínio NudC), encontrada em *Arabidopsis thaliana* e responsável por limitar a extensão do domínio meristemático e/ou promover o desenvolvimento dos domínios dos cotilédones (Jurkuta et al., 2009), além do desenvolvimento e funções de termotolerância (Perez et al., 2009); e GRMZM2G020150, codificando o fator de transcrição AP2-EREBP-transcription factor 196. Um fator de transcrição da família APETALA2, encontrado no arroz, é

responsável por regular negativamente a expressão de genes da síntese do amido tipo I (Fu e Xue, 2010). Como relatado, esse fato pode estar correlacionado com a capacidade do grão de se expandir, pois esse evento depende diretamente do teor de amido contido no grão (Singh et al., 2017).

6.4 Análise comparativa dos resultados da Associação de Modelo Linear Misto (MLMA) e do Mapeamento de Herdabilidade Regional (RHM)

A proporção de regiões associadas às características complexas do milho-pipoca encontradas na abordagem RHM foi significativamente maior do que a observada pelo procedimento MLMA, corroborando a hipótese postulada por Shirali et al. (2016), os quais sugeriram que o método RHM tem um desempenho melhor do que o método SSGWAS — método de associação genômica, no qual todos os SNPs são considerados simultaneamente, juntamente com todos os fenótipos dos indivíduos genotipados e não genotipados, na captura de variantes efetivas (Wang et al., 2014).

Resultados semelhantes foram encontrados por Riggio et al. (2013) ao avaliarem a resistência de nematoides e o peso corporal em cordeiros. Os autores observaram que houve diferença no número de associações identificadas por GWAS e RHM, verificando que, dos 263 QTLs detectados por GWAS e dos 298 QTLs detectados por RHM, apenas 135 foram coincidentes. No entanto, a proporção de variância aditiva observada via RHM (31%) foi menor que a detectada por GWAS (39%). Assim, RHM detectou grandes QTLs raros e, como seus sinais foram estendidos para regiões próximas, outros QTLs menores foram detectados por acaso.

Nagamine et al. (2012) identificaram que, em duas regiões exploradas, nenhum SNP único foi detectado e que o ajuste do SNP mais significativo teve um efeito limitado na herdabilidade regional estimada, sugerindo que essas regiões contêm uma ou mais variantes causais, que não podem ser explicadas pela associação com apenas um SNP, mas podem ser detectadas pela herdabilidade regional.

Para Resende et al. (2017), é esperado que nem todos os locais detectados pelo procedimento RHM sejam observados no GWAS. No entanto, o oposto não deveria ocorrer, a não ser que correspondam aos falsos positivos do GWAS. Para os autores, a contribuição de muitos marcadores de pequeno efeito

pode alcançar significância na fração determinada por RHM; no entanto, um grande efeito de SNP pode não ser significativo na análise via GWAS.

Na análise aqui realizada, verifica-se que a maioria das associações observadas tanto na abordagem MLMA quanto na RHM não coincidiram. Das sete associações constadas via MLMA (Tabela 6), no ENV1, apenas três correspondem a associações detectadas pela análise RHM (Tabela 8), enquanto, no ENV2 (Tabela 7), apenas uma coincidência foi observada na análise RHM para com MLMA (Tabela 9). Segundo postulam Resende et al. (2017), isso pode ocorrer quando uma região com um SNP de grande efeito, detectado pela GWAS, não alcança o nível de significância no RHM, devido a SNPs de pequenos efeitos ajustados no modelo.

Das associações observadas detectadas em ambas as abordagens, o gene GRMZM2G089995, vinculado à característica 100GW, demonstrou estar relacionado com a codificação do fator de transcrição AP2-EREBP-transcription factor 209 (ou fator responsivo do etileno). Os fatores de transcrição AP2/EREBP são muito importantes, pois desempenham várias funções nas plantas (Zhang et al., 2019).

As proteínas APETALA2/etileno (AP2/EREBPs) são os principais reguladores das respostas ao desenvolvimento, crescimento e estresse das plantas (Liu e Zhang, 2017; Zhang et al., 2019). Genes AP2/EREBP desempenham uma função crucial na resposta a vários estresses ambientais em algodão (Liu e Zhang, 2017). Genes AP2/ERF também demonstraram estar diretamente associados a respostas a estresse abiótico em milho, como alagamento (Du et al., 2014) e desidratação e indução a frio em *Arabidopsis* (Sakuma et al., 2002).

Os fatores de transcrição ERF são proteínas AP2 / EREBP que contêm apenas um domínio AP2 e constituem a maior subfamília da família AP2 / EREBP (Zhang et al., 2019).

Não foi identificada, no entanto, uma relação direta com a característica 100GW, com a qual o gene GRMZM2G089995 estaria associado. Sem embargo, verifica-se que o rendimento de grãos é influenciado diretamente pelo estresse abiótico, conforme relatado por Khan et al. (2001), Borrmann et al. (2009), Terra et al. (2013), Zain et al. (2014) e Kamphorst et al. (2018b). É possível sugerir que a característica 100GW está relacionada com respostas a estresses e, quando há

a redução do rendimento de grãos, há também a superexpressão do gene GRMZM2G089995, que codifica AP2/EREBP, relacionada a estresses em plantas.

A proteína contendo o domínio de repetição de tetratricopeptídeo (TPR) é codificada pelo gene GRMZM2G069618, que, por sua vez, demonstrou estar relacionado com a característica GY. A TPR está envolvida com inúmeras funções nos organismos vegetais, entre elas, destacam-se: esterilidade híbrida no arroz (Yu et al., 2016), estresses e sinalização hormonal (Sharma e Pandey, 2016), desenvolvimento da raiz (Zhang et al., 2015b) e falha no desenvolvimento do endosperma, com consequente redução no número de sementes (Awasthi et al., 2012).

O mutante *OsAPC6*, inserido no arroz, está relacionado ao TPR. Foi observado que houve uma redução na produção de sementes de 40 a 45%, o que pode ser atribuído à ausência de endosperma, cuja formação foi anormal em uma alta proporção de gametófitos femininos deste mutante, devido à formação anormal de núcleos polares (Awasthi et al., 2012).

Além disso, a TPR pode estar envolvida com conteúdo de amilose, aparência dos grãos, propriedades físico-químicas (Wu et al., 2015), tamanho dos grãos e qualidade do amido do arroz (She et al., 2010). Os resultados aqui encontrados corroboram as hipóteses postuladas nos trabalhos acima citados, uma vez que a TPR pode estar envolvida diretamente com a expressão da massa de grãos em milho-pipoca, influenciando, conseqüentemente, na característica rendimento de grãos (GY).

O gene GRMZM2G461936, também identificado em ambas as abordagens MLMA e RHM, codifica a proteína Argonaute108 e está associado à característica PE. Proteínas da família Argonaute (AGO) demonstraram estar envolvidas em processos de desenvolvimento de plantas (Lynn et al., 1999; Vaucheret, 2004; Vaucheret et al., 2006; Meng et al., 2013; Zhai et al., 2014). Proteínas AGO regulam o crescimento e o desenvolvimento das plantas com vários pequenos RNAs (sRNAs) (Zhong et al., 2019).

Kidner e Martienssen (2005) relatam que a proteína AGO1 é importante na determinação da identidade do meristema, e, quando da sua deficiência ou mutação, há o comprometimento da formação de folhas e inflorescência em plantas de *Arabidopsis*. No entanto, recentemente, uma proteína denominada

OsAGO17 foi identificada em arroz, regulando positivamente o tamanho e a massa de grãos (Zhong et al., 2019).

Estudos demonstram que a ação da Argonaute está fortemente vinculada à produção de siRNA. Os siRNAs dependentes de Pol IV (p4-) — pequenos RNAs associados ao silenciamento de genes transcricionais, elementos transponíveis e formação de heterocromatina — são mais expressos nos cromossomos maternos no desenvolvimento do endosperma, do que em qualquer outro tecido vegetal de *Arabidopsis* (Mosher et al., 2009; Mosher, 2010; Durán-Figueroa e Vielle-Calzada, 2010).

O que se pode sugerir, com o resultado obtido e com os dados da literatura, é que a proteína Argonaute108, cuja função específica ainda permanece desconhecida, pode estar envolvida não apenas no processo de desenvolvimento em plantas de milho, como também no desenvolvimento do endosperma. O endosperma do milho é composto por dois tipos, sendo eles o translúcido e o opaco, e a proporção de endosperma translúcido parece ser a mais intimamente associada à capacidade de expansão (Hoseney et al., 1983).

O endosperma do milho-pipoca desempenha um papel essencial na capacidade de expansão dos grãos, uma vez que esta é influenciada pela espessura do pericarpo e pelo tipo de amido do endosperma presente no núcleo da semente (Babu et al., 2006).

O gene GRMZM2G110726, associado à característica PV, no ENV2, também demonstrou estar presente em ambas as análises MLMA e RHM. Este gene codifica BOBBER1, que é uma pequena proteína homóloga à NudC (proteína de movimento nuclear). BOBBER1 é uma proteína observada em *Arabidopsis* e descrita como fundamental ao desenvolvimento normal, bem como ao desenvolvimento do meristema após a embriogênese, além da termotolerância em plantas da espécie (Perez et al., 2009). Ainda segundo os autores, mutantes BOB1 apresentaram defeitos no crescimento geral da parte aérea e da raiz, além de defeitos em padrões de desenvolvimento mais específicos.

Além disso, Jurkuta et al. (2009) observaram que a BOBBER1 é necessária para limitar a extensão do domínio meristemático, além de ser responsável por promover o desenvolvimento dos domínios cotiledonares de *Arabidopsis*.

No entanto, não foi possível, neste trabalho, relacionar a função proteica com a característica à qual a proteína BOBBER1 demonstrou estar relacionada. Sugere-se que ela possa estar envolvida no processo de desenvolvimento dos grãos, corroborando os resultados relatados por Perez et al. (2009), considerando-se que essa proteína está envolvida com desenvolvimentos específicos da planta, por conseguinte, não se pode excluir a hipótese de que esteja influenciando a característica volume de pipoca expandida. Porém, devido à falta de informação sobre essa proteína em espécies vegetais, sobretudo no milho, ainda não é possível afirmar sua relação direta com a característica em questão.

CONCLUSÕES

Os resultados revelam uma ampla variação em uma população contendo 196 genótipos diversos, genotipados por 10.507 marcadores SNPs polimórficos e 98 famílias fenotipadas, e mostram que algumas características morfológicas são moderadamente hereditárias, como altura de planta (PH), altura de espiga (EH) e massa de 100 grãos (100GW).

Além disso, foi possível observar que a herdabilidade capturada pelo método MLMA foi, em sua maioria, maior que pelo procedimento RHM. A proporção de herdabilidade capturada em ambas as análises variou de acordo com a característica. Esperava-se que todas as associações encontradas via MLMA também estivessem presentes em RHM, o que não foi verificado, de forma que a posição genômica entre os QTLs da MLMA e do RHM variou de acordo com a característica avaliada.

Apesar do elevado LD, que permitiu um maior tamanho de região genômica, foi possível detectar apenas uma pequena fração dos locos responsáveis pelas principais características em milho-pipoca. Dos SNPs observados em ambas as análises, apenas quatro coincidiram. Não obstante, as associações encontradas, mesmo que em baixo *quantum*, demonstram um valioso recurso para o estudo das características com as quais estão relacionadas, a saber: 100GW, GY, PE e PV. Os genes candidatos subjacentes a estes locos associados fornecem um recurso inestimável para estudar a funcionalidade e dissecar a rede molecular na regulação do desenvolvimento das

características mais importantes do milho-pipoca, sendo sobremaneira útil para a seleção assistida por marcadores destas características em programas de melhoramento.

O uso de linhagens e um maior número de genótipos poderão ser uma alternativa valiosa para a detecção de maior herdabilidade, tanto por MLMA quanto via RHM.

Desde o lançamento do genoma de referência B73, muitas pesquisas foram desenvolvidas utilizando esse genoma e suas anotações na comparação com outros genomas de milho. Contudo, em milho-pipoca, ainda há um marcante déficit de trabalhos que busquem identificar e explicar suas características mais complexas. O trabalho aqui apresentado vem, pois, auxiliar essa identificação, fornecendo resultados importantes sobre a arquitetura genética de características importantes da espécie. Este foi um estudo preliminar e pode ser aprimorado em pesquisas futuras por meio do uso de linhagens e com maior quantidade de genótipos, a fim de se obter melhores resultados e, dessa forma, contribuir ainda mais com a evolução dos programas de melhoramento com essa importante cultura.

REFERÊNCIAS BIBLIOGRÁFICAS

- Akiyama, T., Pillai, M.A., Sentoku, N. (2004) Cloning, characterization and expression of OsGLN2, a rice endo-1,3- β -glucanase gene regulated developmentally in flowers and hormonally in germinating seeds. *Planta*, 220: 129-139.
- Almeida, V.C., Viana, J.M.S., Oliveira, H.M., Risso, L.A., Ribeiro, A.F.S., DeLima, R.O. (2018a) Genetic diversity and path analysis for nitrogen use efficiency of tropical popcorn (*Zea mays* ssp. *everta*) inbred lines in adult stage. *Plant Breed* 137:839–847.
- Almeida, V.C., Viana, J.M.S., Risso, L.A., Ribeiro, C., Lima, R.O. (2018b) Generation mean analysis for nitrogen and phosphorus uptake, utilization, and translocation indexes at vegetative stage in tropical popcorn. *Euphytica* 214(7):103.
- Alves, M.F., Nizio, D.A.C., Brito, F.A., Sampaio, T.S., Silva, A.V.C., Arrigoni-Blank, M.F., Carvalho, S.V.A., Blank, A.F. (2016) Analysis of genetic diversity of a native population of *Myrcia lundiana* Kiaersk. plants using ISSR markers. *Genet. Mol. Res.* 15:(4).

- Alves, M.L., Carbas, B., Gaspar, D., Paulo, M., Brites, C., Mendes-Moreira, P., Brites, C.M., Malosetti, M., van Eeuwijk, F., Vaz Patto, M.C. (2019) Genome-wide association study for kernel composition and flour pasting behavior in wholemeal maize flour. *BMC Plant Biol* 19(1):123.
- Amaral Junior, A.T., Freitas Junior, S.P., Rangel, R.M., Pena, G.F., Ribeiro, R.M., Morais, R.C., Schuelter, A.R. (2010) Improvement of a popcorn population using selection indexes from a fourth cycle of recurrent selection program carried out in two different environments. *Genet. Mol. Res.* 9(1):340-347.
- Amaral Junior, A.T., Freitas, I.L. de J., Guimarães, A.G., Maldonado, C., Arriagada, O., Mora, F. (2016) Bayesian analysis of quantitative traits in popcorn (*Zea mays* L.) through four cycles of recurrent selection. *Plant Prod. Sci.* 19(4): 574-578.
- Amaral Júnior, A.T., Gonçalves, L.S.A., Freitas Júnior, S. de P., Candido, L.S., Vittorazzi, C., Pena, G.F., Ribeiro, R.M., Silva, T.R. da C., Pereira, M.G., Scapim, C.A., Viana, A.P., Carvalho, G.F. (2013) UENF 14: a new popcorn cultivar. *Crop Breeding and Applied Biotechnology*, 13: 218-220.
- Amaral Júnior, A.T., Poltronieri, T.P.S., Santos, P.H.D., Vivas, M., Gerhardt, I.F.S., Carvalho, B.M., Freitas, C.S., Silveira, S.F. (2019) Reaction of popcorn lines (s7) cultivated in distinct phosphorus levels to *Bipolaris maydis* infection. *Summa Phytopathol.* 45:18-22.
- Anderson, S.N., Stitzer, M.C., Brohammer, A.B., Zhou, P., Noshay, J.M., O'Connor, C.H., Hirsch, C.D., Ross- Ibarra, J., Hirsch, C.N., Springer, N.M. (2019) Transposable elements contribute to dynamic genome content in maize. *Plant J.* 100:1052-1065.
- Andorf, C.M., Lawrence, C.J., Harper, L.C., Schaeffer, M.L., Campbell, D.A., Sen, T.Z. (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.

- Andrade, A.C.B., Viana, J.M.S., Pereira, H.D., Pinto, V.B., Fonseca e Silva, F. (2019) Linkage disequilibrium and haplotype block patterns in popcorn populations. *PLoS One* 14:e0219417.
- Arnhold, E., Mora, F., G. Silva, R., I.V. Good-God, P., Rodovalho, M.A. (2009) Evaluation of top-cross popcorn hybrids using Mixed Linear Model methodology. *Chilean Journal of Agricultural Research*, 69(1):46-53.
- Arthikala, M.K., Montiel, J., Sánchez-López, R., Nava, N., Cárdenas, L., Quinto, C. (2017) Respiratory Burst Oxidase Homolog Gene A Is Crucial for Rhizobium Infection and Nodule Maturation and Function in Common Bean. *Front Plant Sci.* 10.3389/fpls.2017.02003
- Arthikala, M.K., Quinto, C. (2018) RbohA coordinates lateral root emergence in common bean. *Commun. Integr. Biol.* 11:1-5.
- Avisar, D., Abu-Abied, M., Belausov, E., Sadot, E., Hawes, C., Sparkes, I.A. (2009) A Comparative Study of the Involvement of 17 *Arabidopsis* Myosin Family Members on the Motility of Golgi and Other Organelles. *Plant. Physiol.*150:700-709.
- Awasthi, A., Paul, P., Kumar, S., Verma, S.K., Prasad, R., Dhaliwal, H.S. (2012) Abnormal endosperm development causes female sterility in rice insertional mutant *OsAPC6*. *Plant. Sci.* 183:167-174.
- Babu, R., Nair, S.K., Kumar, A., Rao, H.S., Verma, P., Gahalain, A., Singh, I.S., Gupta, H.S. (2006) Mapping QTLs for popping ability in a popcorn × flint corn cross. *Theor. Appl. Genet.* 112:1392-1399.
- Basirnia, A., Hatami Maleki, H., Darvishzadeh, R., Ghavami, F. (2014) Mixed linear model association mapping for low chloride accumulation rate in oriental-type tobacco (*Nicotiana tabaccum* L.) germplasm. *J. Plant Interact* 9: 666-672.
- Bates, D., Mächler, M., Bolker, B., Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 10.18637/jss.v067.i01

- Bilinski, P., Albert, P.S., Berg, J.J., Birchler, J.A., Grote, M.N., Lorant, A., Quezada, J., Swarts, K., Yang, J., Ross-Ibarra, J. (2018) Parallel altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*. *PLOS Genet.* 14: e1007162.
- Bloom, A.J., Meyerhoff, P.A., Taylor, A.R., Rost, T.L. (2002) Root Development and Absorption of Ammonium and Nitrate from the Rhizosphere. *J. Plant Growth Regul* 21:416-431.
- Bolduc, N., Yilmaz, A., Mejia-Guerra, M.K., Morohashi, K., O'Connor, D., Grotewold, E., Hake, S. (2012) Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes Dev.* 26:1685-1690.
- Borrmann, D., Junqueira, R. de M., Sinnecker, P., Gomes, M.S. de O., Castro, I.A., Marquez, U.M.L. (2009) Chemical and biochemical characterization of soybean produced under drought stress. *Ciê. e Tecnol. Aliment.* 29:676-681.
- Brachi, B., Morris, G.P., Borevitz, J.O. (2011) Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol.* 12:232.
- Brohammer, A.B., Kono, T.J.Y., Springer, N.M., McGaugh, S.E., Hirsch, C.N. (2018) The limited role of differential fractionation in genome content variation and function in maize (*Zea mays* L.) inbred lines. *Plant J.* 93:131-141.
- Caballero, A., Tenesa, A., Keightley, P.D. (2015) The Nature of Genetic Variation for Complex Traits Revealed by GWAS and Regional Heritability Mapping Analyses. *Genetics* 201:1601-1613.
- Capovilla, G., Pajoro, A., Immink, R.G., Schmid, M. (2015) Role of alternative pre-mRNA splicing in temperature signaling. *Curr. Opin. Plant Biol.* 27:97-103.
- Carpentieri-Pípolo, V., Rinaldi, D.A., Lima, V.E.N. de. (2005) Adaptabilidade e estabilidade de populações de milho-pipoca. *Pesqui. Agropecuária Bras.* 40:87-90.
- Carpentieri-Pípolo, V., Silva, F.A.M., Seifert, A.L. (2003) Popcorn parental selection based on genetic divergence. *Crop Breeding and Applied Biotechnology*, 3:261-268.

- Chai, W., Jiang, P., Huang, G., Jiang, H., Li, X. (2017) Identification and expression profiling analysis of TCP family genes involved in growth and development in maize. *Physiol. Mol. Biol. Plants* 23:779-791.
- Chen, J., Shrestha, R., Ding, J., Zheng, H., Mu, C., Wu, J., Mahuku, G. (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.
- Chen, L., Li, Y., Li, C., Shi, Y., Song, Y., Zhang, D., Li, Y., Wang, T. (2018) Genome-wide analysis of the pentatricopeptide repeat gene family in different maize genomes and its important role in kernel development. *BMC Plant. Biol.* 18:366.
- Cheng, L.B., Li, S.Y., Yang, G.X., Jing, X.M., He, G.Y., Mones, N.G. (2010) Overexpression of Soybean (*Glycine max* (L.) Meer.) L34 Gene Leads to Reduced Survival to Cold Stress in Transgenic Arabidopsis. *Plant Mol Biol Report* 28:41-48.
- Cheng, X., Zhu, X., Tian, W., Cheng, W., Hakim., Sun, J., Jin, S., Zhu, H. (2017) Genome-wide identification and expression analysis of polyamine oxidase genes in upland cotton (*Gossypium hirsutum* L.). *Plant Cell, Tissue and Organ Cult.* 129:237-249.
- Chia, J., Song, C., Bradbury, P.J., Costich, D., de Leon, N., Doebley, J., Elshire, R.J., Gaut, B., Geller, L., Glaubitz, J.C., Gore, M., Guill, K.E., Holland, J., Hufford, M.B., Lai, J., Li, M., Liu, X., Lu, Y., McCombie, R., Nelson, R., Poland, J., Prasanna, B.M., Pyhäjärvi, T., Rong, T., Sekhon, R.S., Sun, Q., Tenailon, M.I., Tian, F., Wang, J., Xu, X., Zhang, Z., Kaeppler, S.M., Ross-Ibarra, J., McMullen, M.D., Buckler, E.S., Zhang, G., Xu, Y., Ware, D. (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* 44:803-807.
- Clifford, D., McCullagh, P. (2012) The regress package. R Package version. 1-3.

- Climate-data.org. (2020) Clima: Campos dos Goytacazes: <https://pt.climate-data.org/america-do-sul/brasil/rio-de-janeiro/campos-dos-goytacazes-4053/> em 23/08/2020.
- Coan, M.M.D., Senhorinho, H.J.C., Pinto, R.J.B., Scapim, C.A., Tessmann, D.J., Williams, W.P., Warburton, M.L. (2018a) Genome-Wide Association Study of Resistance to Ear Rot by in a Tropical Field Maize and Popcorn Core Collection. *Crop. Sci.* 58:564.
- Coan, M.M.D., Senhorinho, H.J.C., Pinto, R.J.B., Scapim, C.A., Tessmann, D.J., Williams, W.P., Warburton, M.L. (2018b) Genome-Wide Association Study of Resistance to Ear Rot by *Fusarium verticillioides* in a Tropical Field Maize and Popcorn Core Collection. *Crop Science*, 58(2):564-578.
- Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B., Pang, E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142:169-196.
- Comstock, R.E., Robinson, H.F. (1948) The Components of Genetic Variance in Populations of Biparental Progenies and Their Use in Estimating the Average Degree of Dominance. *Biometrics*, 4:254.
- Daros, M., Amaral Junior, A.T. do., Pereira, M.G., Santos, F.S., Gabriel, A.P.C., Scapim, C.A., Freitas Jr., S. de P., Silvério, L. (2004) Recurrent selection in inbred popcorn families. *Sci. Agric.* 61:609-614.
- Daros, M., Amaral Júnior, A.T., Pereira, M.G. (2002) Genetic gain for grain yield and popping expansion in full-sib recurrent selection in popcorn. *Crop Breeding and Applied Biotechnology*, 2:339-344.
- Dietz, K., Vogel, M.O., Viehhauser, A. (2010) AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. *Protoplasma*, 245:3-14.

- Díez, C.M., Gaut, B.S., Meca, E., Scheinvar, E., Montes-Hernandez, S., Eguiarte, L.E., Tenaillon, M.I. (2013) Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytol.* 199:264-276.
- Donnini, S., Prinsi, B., Negri, A.S., Vigani, G., Espen, L., Zocchi, G. (2010) Proteomic characterization of iron deficiency responses in *Cucumis sativus* L. roots. *BMC Plant Biol.* 10:268.
- Doyle, J.J., Doyle, J.L. (1990) Isolation of plant DNA from fresh tissue. *Focus (Madison)* 12:13-15.
- Du, H., Huang, M., Zhang, Z., Cheng, S. (2014) Genome-wide analysis of the AP2/ERF gene family in maize waterlogging stress response. *Euphytica* 198:115-126.
- Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., Abraham, C., Regueiro, M., Griffiths, A., Dassopoulos, T., Bitton, A., Yang, H., Targan, S., Datta, L.W., Kistner, E.O., Schumm, L.P., Lee, A.T., Gregersen, P.K., Barmada, M.M., Rotter, J.I., Nicolae, D.L., Cho, J.H. (2006) A Genome-Wide Association Study Identifies IL23R as an Inflammatory Bowel Disease Gene. *Science*, 314(5804):1461-1463.
- Durán-Figueroa, N., Vielle-Calzada, J.-P. (2010) ARGONAUTE9-dependent silencing of transposable elements in pericentromeric regions of *Arabidopsis*. *Plant Signal. Behav.* 5:1476-1479.
- Dutt, S., Parkash, J., Mehra, R., Sharma, N., Singh, B., Raigond, P., Joshi, A., Chopra, S., Singh, B.P. (2015) Translation initiation in plants: roles and implications beyond protein synthesis. *Biol. Plant.* 59:401-412.
- Earl, D.A., VonHoldt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359-361.
- Endelman, J.B. (2011) Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *Plant Genome J.* 4:250.

- Ertiro, B.T., Semagn, K., Das, B., Olsen, M., Labuschagne, M., Worku, M., Wegary, D., Azmach, G., Ogugo, V., Keno, T., Abebe, B., Chibsa, T., Menkir, A. (2017) Genetic variation and population structure of maize inbred lines adapted to the mid-altitude sub-humid maize agro-ecology of Ethiopia using single nucleotide polymorphic (SNP) markers. *BMC Genomics*, 18:777.
- Evanno, G., Regnaut, S., Goudet, J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14:2611-2620.
- Falconer, D.D. (1981) *Introdução a genética quantitativa*. UFV. Viçosa 279p.
- Foreman, J., Demidchik, V., Bothwell, J.H.F., Mylona, P., Miedema, H., Torres, M.A., Linstead, P., Costa, S., Brownlee, C., Jones, J.D.G., Davies, J.M., Dolan, L. (2003) Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. *Nature*, 422:442-446.
- Frascaroli, E., Schrag, T.A., Melchinger, A.E. (2013) Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. *Theor. Appl. Genet.* 126:133-141.
- Freitas, I.L.J., do Amaral Júnior, A.T., Freitas Jr., S.P., Cabral, P.D.S., Ribeiro, R.M., Gonçalves, L.S.A. (2014) Genetic gains in the UENF-14 popcorn population with recurrent selection. *Genet. Mol. Res.* 13:518-527.
- Freitas Júnior, S.P., Amaral Junior, A.T., Rangel, R.M., Viana, A. P. (2009) Genetic gains in popcorn by full-sib recurrent selection. *Crop Breeding and Applied Biotechnology*, 9:1-7.
- Fu, F., Xue, H. (2010) Coexpression Analysis Identifies Rice Starch Regulator1, a Rice AP2/EREBP Family Transcription Factor, as a Novel Rice Starch Biosynthesis Regulator. *Plant Physiol.* 154:927-938.

- Gage, J.L., Vaillancourt, B., Hamilton, J.P., Manrique-Carpintero, N.C., Gustafson, T.J., Barry, K., Lipzen, A., Tracy, W.F., Mikel, M.A., Kaeppler, S.M., Buell, C.R., de Leon, N. (2019) Multiple Maize Reference Genomes Impact the Identification of Variants by Genome-Wide Association Study in a Diverse Inbred Panel. *Plant. Genome*, 10.3835/plantgenome2018.09.0069.
- Ganal, M.W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E.S., Charcosset, A., Clarke, J.D., Graner, E.-M., Hansen, M., Joets, J., Le Paslier, M.C., McMullen, M.D., Montalent, P., Rose, M., Schön, C.C., Sun, Q., Walter, H., Martin, O.C., Falque, M. (2011) A Large Maize (*Zea mays* L.) SNP Genotyping Array: Development and Germplasm Genotyping, and Genetic Mapping to Compare with the B73 Reference Genome. *PLoS One*, 6: e28334.
- Ge, W., Song, Y., Zhang, C., Zhang, Y., Burlingame, A.L., Guo, Y. (2011) Proteomic analyses of apoplastic proteins from germinating *Arabidopsis thaliana* pollen. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1814:1964-1973.
- Geldermann, H. (1975) Investigations on inheritance of quantitative characters in animals by gene markers I. Methods. *Theor. Appl. Genet* 46:319-330.
- Gerhardt, I.F.S., Júnior, A.T.A., Guimarães, L.J.M., Schwantes, I.A., Santos, A., Kamphorst, S.H., Lima, V.J., Poblete, F.M., Myers, G.O. (2017a) Intraspecific variability of popcorn S7 lines for phosphorus efficiency in the soil. *Genet. Mol. Res.* 10.4238/gmr16039763
- Gerhardt, I.F.S., Amaral Junior, A.T., Pena, G.F, Guimarães, L.J.M., Lima, V.J., Vivas, M., Santos, P.H.A.D., Ferreira, F.R.A., Freitas, M.S.M., Kamphorst, S.H. (2019) Genetic effects on the efficiency and responsiveness to phosphorus use in popcorn as estimated by diallel analysis. *PLoS One*, 14(5).
- Gu, R., Duan, F., An, X., Zhang, F., von Wirén, N., Yuan, L. (2013) Characterization of AMT-Mediated High-Affinity Ammonium Uptake in Roots of Maize (*Zea mays* L.). *Plant Cell Physiol.* 54:1515-1524.

- Guan, L., Cao, K., Li, Y., Guo, J., Xu, Q., Wang, L. (2019) Detection and application of genome-wide variations in peach for association and genetic relationship analysis. *BMC Genet.* 20:101.
- Guimarães, A.G., Amaral Júnior, A.T., Almeida Filho, J.E., Pena, G.F., Vittorazzi, C., Pereira, M.G. (2018a) Population structure and impact of recurrent selection on popcorn using EST-SSR markers. *Acta Sci. Agron.* 40: e35218.
- Guimarães, A.G., Amaral Júnior, A.T., Lima, V.J., Leite, J.T., Scapim, C.A., Vivas, M. (2018b) Genetic gains and selection advances of the UENF-14 popcorn population. *Rev. Caatinga*, 31:271-278.
- Guimarães, A.G., Amaral Júnior, A.T., Pena, G.F., Almeida filho, J.E., Pereira, M.G., Santos, P.H.A.D. (2019) Genetic gains in the popcorn population UENF-14: developing the ninth generation of intrapopulation recurrent selection. *Rev Caatinga*, 32:625-633.
- Haiman, C.A., Patterson, N., Freedman, M.L., Myers, S.R., Pike, M.C., Waliszewska, A., Neubauer, J., Tandon, A., Schirmer, C., McDonald, G.J., Greenway, S.C., Stram, D.O., Le Marchand, L., Kolonel, L.N., Frasco, M., Wong, D., Pooler, L.C., Ardlie, K., Oakley-Girvan, I., Whittemore, A.S., Cooney, K.A., John, E.M., Ingles, S.A., Altshuler, D., Henderson, B.E., Reich, D. (2007) Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat. Genet.* 39:638-644.
- Han, B., Huang, X. (2013) Sequencing-based genome-wide association study in rice. *Curr. Opin. Plant. Biol.* 16:133-138.
- Han, S., Miedaner, T., Utz, H.F., Schipprack, W., Schrag, T.A., Melchinger, A.E. (2018) Genomic prediction and GWAS of Gibberella ear rot resistance traits in dent and flint lines of a public maize breeding program. *Euphytica*, 214:6.
- Hauser, F., Strassner, J., Schaller, A. (2001) Cloning, expression, and characterization of tomato (*Lycopersicon esculentum*) aminopeptidase P. *J. Biol. Chem.* 276:31732-31737.
- Hazel, L.N. (1943) The genetic basis for constructing selection indexes. *Genetics*, 28: 476–90.

- Hill, W.G., Weir, B.S. (1988) Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* 33:54-78.
- Hirsch, C.N., Hirsch, C.D., Brohammer, A.B., Bowman, M.J., Soifer, I., Barad, O., Shem-Tov, D., Baruch, K., Lu, F., Hernandez, A.G., Fields, C.J., Wright, C.L., Koehler, K., Springer, N.M., Buckler, E., Buell, C.R., de Leon, N., Kaeppeler, S.M., Childs, K.L., Mikel, M.A. (2016) Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. *Plant Cell.* 28:2700-2714.
- Hoffmeister, M., Piotrowski, M., Nowitzki, U., Martin, W. (2005) Mitochondrial trans-2-Enoyl-CoA reductase of wax ester fermentation from *Euglena gracilis* defines a new family of enzymes involved in lipid synthesis. *J. Biol. Chem.* 280:4329-4338.
- Hoseney, R.C., Zeleznak, K., Abdelrahman, A. (1983) Mechanism of popcorn popping. *J. Cereal. Sci.* 1:43-52.
- Howitt, S.M., Udvardi, M.K. (2000) Structure, function and regulation of ammonium transporters in plants. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1465:152-170.
- Hsu, K., Liu, C., Wu, S., Kuo, Y., Lu, C., Wu, C., Lian, P., Hong, C., Ke, Y., Huang, J., Yeh, C. (2014) Expression of a gene encoding a rice RING zinc-finger protein, OsRZFP34, enhances stomata opening. *Plant. Mol. Biol.* 86:125-137.
- Hu, L., Li, Y., Xu, W., Zhang, Q., Zhang, L., Qi, X., Dong, H. (2012) Improvement of the photosynthetic characteristics of transgenic wheat plants by transformation with the maize C4 phosphoenolpyruvate carboxylase gene. *Plant. Breed.* 131:385-391.
- Huang, W., Mackay, T.F.C. (2016) The Genetic Architecture of Quantitative Traits Cannot Be Inferred from Variance Component Analysis. *PLOS Genet* 12:e1006421.

- Souza, S.G.H., Carpentieri-Pípolo, V., Garbúglio, D.D., Júnior, N. da S.F., Ruas, C. de F., Ruas, P.M. (2012) Genetic distance estimated by RAPD markers and performance of topcross hybrids in popcorn. *Am. J. Plant Scien.* 3:1666-1673.
- Iglesias-Fernández, R., Rodríguez-Gacio, M.C., Barrero-Sicilia, C., Carbonero, P., Matilla, A. (2011) Three endo- β -mannanase genes expressed in the micropylar endosperm and in the radicle influence germination of *Arabidopsis thaliana* seeds. *Planta*, 233:25-36.
- Instituto Agronômico de Campinas. (2019) Instituto Agronômico de Campinas – IAC: <http://www.iac.sp.gov.br/produtoseservicos/sementesiacy/> em 20/10/19.
- Jamann, T.M., Luo, X., Morales, L., Kolkman, J.M., Chung, C.-L., Nelson, R.J. (2016) A remorin gene is implicated in quantitative disease resistance in maize. *Theor. Appl. Genet.* 129:591-602.
- Jia, F., Wan, X., Zhu, W., Sun, D., Zheng, C., Liu, P., Huang, J. (2015) Overexpression of Mitochondrial Phosphate Transporter 3 Severely Hampers Plant Development through Regulating Mitochondrial Function in *Arabidopsis*. *PLoS One*, 10:e0129717.
- Jian, Y., Xu, C., Guo, Z., Wang, S., Xu, Y., Zou, C. (2017) Maize (*Zea mays* L.) genome size indicated by 180-bp knob abundance is associated with flowering time. *Sci. Rep.* 7:5954.
- Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M.C., Wang, B., Campbell, M.S., Stein, J.C., Wei, X., Chin, C.-S., Guill, K., Regulski, M., Kumari, S., Olson, A., Gent, J., Schneider, K.L., Wolfgruber, T.K., May, M.R., Springer, N.M., Antoniou, E., McCombie, W.R., Presting, G.G., McMullen, M., Ross-Ibarra, J., Dawe, R.K., Hastie, A., Rank, D.R., Ware, D. (2017) Improved maize reference genome with single-molecule technologies. *Nature*, 546:524-527.
- Jue, D., Sang, X., Lu, S., Dong, C., Zhao, Q., Chen, H., Jia, L. (2015) Genome-Wide Identification, Phylogenetic and Expression Analyses of the Ubiquitin-Conjugating Enzyme Gene Family in Maize. *PLoS One*, 10:e0143488.

- Jurkuta, R.J., Kaplinsky, N.J., Spindel, J.E., Barton, M.K. (2009) Partitioning the Apical Domain of the *Arabidopsis* Embryo Requires the BOBBER1 NudC Domain Protein. *Plant Cell*, 21:1957-1971.
- Kamphorst, S.H., Amaral Júnior, A.T., Lima, V.J., Guimarães, L.J.M., Schmitt, K.F.M., Leite, J.T., Santos, P.H.A.D., Chaves, M.M., Mafra, G.S., Santos Junior, D.R. dos., Cruz, C.D., Campostrini, E. (2019) Can Genetic Progress for Drought Tolerance in Popcorn Be Achieved by Indirect Selection? *Agronomy*, 9:792.
- Kamphorst, S.H., Lima, V.J., Amaral Júnior, A.T.A., Schmitt, K.F.M., Leite, J.T., Carvalho, C.M., Silva, R.M.R., Xavier, K.B., Ferreira, F.R.A., Santos, P.H.A., Campostrini, E. (2018a) Popcorn breeding for water-stress tolerance or for agronomic water-use efficiency? *Genet. Mol. Res.* 10.4238/gmr18184
- Kamphorst, S.H., Lima, V.J., Schmitt, K.F.M., Leite, J.T., Azeredo, V.C., Pena, G.F., Santos, P.H.A.D., Júnior, D.R.S., Júnior, S.B. da S., Bispo, R.B., Santos, T.O., Khan, S., Júnior, A.T. do A. (2018b) Water stress adaptation of popcorn roots and association with agronomic traits. *Genet. Mol. Res.* 17(3).
- Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S., Freimer, N.B., Sabatti, C., Eskin, E. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42:348-354.
- Kermgard, E., Yang, Z., Michel, A.-M., Simari, R., Wong, J., Ibba, M., Lazazzera, B.A. (2017) Quality Control by Isoleucyl-tRNA Synthetase of *Bacillus subtilis* Is Required for Efficient Sporulation. *Sci. Rep.* 7:41763.
- Khan, M.B., Hussain, N., Iqbal, M. (2001) Effect of water stress on growth and yield components of maize variety YHS 202. *J. Res.* 12:15-18.
- Kidner, C.A., Martienssen, R.A. (2005) The role of ARGONAUTE1 (AGO1) in meristem formation and identity. *Dev. Biol.* 280:504-517.
- Kim, M., Hersh, L.B., Leissring, M.A., Ingelsson, M., Matsui, T., Farris, W., Lu, A., Hyman, B.T., Selkoe, D.J., Bertram, L., Tanzi, R.E. (2007) Decreased Catalytic Activity of the Insulin-degrading Enzyme in Chromosome 10-Linked Alzheimer Disease Families. *J. Biol. Chem.* 282:7825-7832.

- Kong, F., Deng, Y., Zhou, B., Wang, G., Wang, Y., Meng, Q. (2014) A chloroplast-targeted DnaJ protein contributes to maintenance of photosystem II under chilling stress. *J. Exp. Bot.* 65:143–158.
- Korte, A., Vilhjálmsson, B.J., Segura, V., Platt, A., Long, Q., Nordborg, M. (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat. Genet.* 44:1066-1071.
- Kotera, M., Tokimatsu, T., Nakagawa, Z.I., Moriya, Y., Hattori, M., Goto, S., Kanehisa, M. (2009) Reclassification of oxidoreductase subclasses based on the relationships of reductants and oxidants. *In: 20th International Conference on Genome Informatics.*
- Krug, C.A., Viégas, G.P., Paoliéri, L. (1943) Híbridos comerciais de milho. *Bragantia*, 3:367-552.
- Kuki, M.C., Scapim, C.A., Rossi, E.S., Mangolin, C.A., Amaral Júnior, A.T. do., Pinto, R.J.B. (2018) Genome wide association study for gray leaf spot resistance in tropical maize core. *PLoS One*, 13:e0199539.
- Kuramata, M., Masuya, S., Takahashi, Y., Kitagawa, E., Inoue, C., Ishikawa, S., Youssefian, S., Kusano, T. (2009) Novel Cysteine-Rich Peptides from *Digitaria ciliaris* and *Oryza sativa* Enhance Tolerance to Cadmium by Limiting its Cellular Accumulation. *Plant Cell Physiol.* 50:106-117.
- Kurosawa, R.N.F., Fernando, I., Gerhardt, S., Neto, R.F. (2017a) GGE Biplot projection in discriminating the efficiency of popcorn lines to use nitrogen. *Ciência e Agrotecnologia*, 41(1):22-3.
- Kurosawa, R.N.F., Vivas, M., Amaral Junior, A.T., Ribeiro, R.M., Miranda, S.B., Pena, G.F., Leite, J.T., Mora, F. (2017b) Popcorn germplasm resistance to fungal diseases caused by *Exserohilum turcicum* and *Bipolaris maydis*. *Bragantia*, 77:36-47.
- Lander, E.S., Botstein, D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121:185-99.

- Larish, L.L.B., Brewbaker, J.L. (1999) Diallel analyses of temperate and tropical popcorns. *Maydica*, 44:279-284.
- Lawrence, C.J. (2004) MaizeGDB, the community database for maize genetics and genomics. *Nucleic. Acids Res.* 32:393-397.
- Lee, K.-J., Kim, K. (2015) The rice serine/threonine protein kinase OsPBL1 (ORYZA SATIVA ARABIDOPSIS PBS1-LIKE 1) is potentially involved in resistance to rice stripe disease. *Plant Growth Regul.* 77:67-75.
- Lenth, R. V. (2016) Least-Squares Means: The R Package lsmeans. *J. Stat Softw.*10.18637/jss.v069.i01
- Li, C., Wang, M., Ma, X., Zhang, W. (2014) NRGA1, a Putative Mitochondrial Pyruvate Carrier, Mediates ABA Regulation of Guard Cell Ion Channels and Drought Stress Responses in *Arabidopsis*. *Mol. Plant.* 7:1508-1521.
- Li, C., Xiang, X., Huang, Y., Zhou, Y., An, D., Dong, J., Zhao, C., Liu, H., Li, Y., Wang, Q., Du, C., Messing, J., Larkins, B.A., Wu, Y., Wang, W. (2020) Long-read sequencing reveals genomic structural variations that underlie creation of quality protein maize. *Nat. Commun.* 11:17.
- Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y. (2012) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45:43-50.
- Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., Liu, J., Warburton, M.L., Cheng, Y., Hao, X., Zhang, P., Zhao, J., Liu, Y., Wang, G., Li, J., Yan, J. (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45:43-50.
- Li, J.Z., Zhang, Z.W., Li, Y.L., Wang, Q.L., Zhou, Y.G. (2011) QTL consistency and meta-analysis for grain yield components in three generations in maize. *Theor. Appl. Genet.* 122:771-782.
- Li, Q., Li, J., Liu, S., Huang, J., Lin, H., Wang, K., Cheng, X., Liu, Z. (2015) A Comparative Proteomic Analysis of the Buds and the Young Expanding

- Leaves of the Tea Plant (*Camellia sinensis* L.). *Int. J. Mol. Sci.* 16:14007-14038.
- Li, Y., Dong, Y., Niu, S., Cui, D. (2007) The genetic relationship among plant-height traits found using multiple-trait QTL mapping of a dent corn and popcorn cross. *Genome*, 50:357-364.
- Li, Y., Kabbage, M., Liu, W., Dickman, M.B. (2016) Aspartyl Protease-Mediated Cleavage of BAG6 Is Necessary for Autophagy and Fungal Resistance in Plants. *Plant Cell*. 28:233-247.
- Lima, V.J., Amaral Júnior, A.T., Kamphorst, S.H., Bispo, R.B., Leite, J.T., Santos, T. de O., Schmitt, K.F.M., Chaves, M.M., Oliveira, U.A., Santos, P.H.A.D., Gonçalves, G.M.B., Khan, S., Guimarães, L.J.M. (2019) Combined Dominance and Additive Gene Effects in Trait Inheritance of Drought-Stressed and Full Irrigated Popcorn. *Agronomy*, 9:782.
- Lima, V.J., Freitas Junior, S.P., Souza, Y.P., Silva, C.S., Farias, J.E.C., Souza, R.F., Chaves, M.M., Feitosa, J.V. (2018) Genetic gain capitalization in the first cycle of recurrent selection in popcorn at Ceará's Cariri. *Rev. Bras. Ciências Agrárias - Brazilian J. Agric. Sci.* 13:1-7.
- Liu, C., Zhang, T. (2017) Expansion and stress responses of the AP2/EREBP superfamily in cotton. *BMC Genomics*, 18:118.
- Liu, N., Xue, Y., Guo, Z., Li, W., Tang, J. (2016) Genome-Wide Association Study Identifies Candidate Genes for Starch Content Regulation in Maize Kernels. *Front. Plant. Sci.* 10.3389/fpls.2016.01046
- Liu, X., Fu, J., Gu, D., Liu, W., Liu, T., Peng, Y., Wang, J., Wang, G. (2008) Genome-wide analysis of gene expression profiles during the kernel development of maize (*Zea mays* L.). *Genomics*, 91:378-387.
- Lovegrove, A., Wilkinson, M.D., Freeman, J., Pellny, T.K., Tosi, P., Saulnier, L., Shewry, P.R., Mitchell, R.A.C. (2013) RNA Interference Suppression of Genes in Glycosyl Transferase Families 43 and 47 in Wheat Starchy Endosperm Causes Large Decreases in Arabinoxylan Content. *Plant Physiol.* 163:95-107.

- Lu, F., Romay, M.C., Glaubitz, J.C., Bradbury, P.J., Elshire, R.J., Wang, T., Li, Y., Li, Y., Semagn, K., Zhang, X., Hernandez, A.G., Mikel, M.A., Soifer, I., Barad, O., Buckler, E.S. (2015) High-resolution genetic mapping of maize pan-genome sequence anchors. *Nat. Commun.* 6:6914.
- Luo, M., Zhao, Y., Zhang, R., Xing, J., Duan, M., Li, J., Wang, N., Wang, W., Zhang, S., Chen, Z., Zhang, H., Shi, Z., Song, W., Zhao, J. (2017) Mapping of a major QTL for salt tolerance of mature field-grown maize plants based on SNP markers. *BMC Plant Biol.* 17:140.
- Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., Barton, M.K. (1999) The PINHEAD/ZWILLE gene acts pleiotropically in *Arabidopsis* development and has overlapping functions with the ARGONAUTE1 gene. *Development*, 126(3):469-481.
- Mafra, G.S., Amaral Junior, A.T., Vivas, M., Santos, J.S., Silva, F.H.L., Guimarães, A.G., Pena, G.F. (2018) The combining ability of popcorn S7 lines for *Puccinia polysora* resistance purposes. *Bragantia*, 77:519-526.
- Magomedova, L., Tiefenbach, J., Zilberman, E., Le Billan, F., Voisin, V., Saikali, M., Boivin, V., Robitaille, M., Gueroussov, S., Irimia, M., Ray, D., Patel, R., Xu, C., Jeyasuria, P., Bader, G.D., Hughes, T.R., Morris, Q.D., Scott, M.S., Krause, H., Angers, S., Blencowe, B.J., Cummins, C.L. (2019) ARGLU1 is a transcriptional coactivator and splicing regulator important for stress hormone signaling and development. *Nucleic Acids Res.* 47:2856-2870.
- Maize Genetics and Genome Database. (2019) Maize Genetics and Genome Database: <https://www.maizegdb.org/> em 11/10/19.
- Maldonado, C., Mora, F., Scapim, C.A., Coan, M. (2019) Genome-wide haplotype-based association analysis of key traits of plant lodging and architecture of maize identifies major determinants for leaf angle: hapLA4. *PLoS One*, 14:e0212925.
- Manchanda, N., Andorf, C.M., Ye, L., Wimalanathan, K., Rounsley, S., Wang, K., Lawrence-Dill, C. (2016) Sequencing, assembly, and annotation of Maize B104 : A maize transformation resource. *58th Annu Maize Genet Conf* 92.

- Manohar, M., Choi, H.W., Manosalva, P., Austin, C.A., Peters, J.E., Klessig, D.F. (2017) Plant and Human MORC Proteins Have DNA-Modifying Activities Similar to Type II Topoisomerases, but Require One or More Additional Factors for Full Activity. *Mol. Plant-Microbe Interact.* 30:87-100.
- Martins, T.G., Freitas Júnior, S.P., Luz, L.N., Marco, C.A., Vásquez, E.M.F. (2018) Inoculation efficiency of *Azospirillum brasilense* on economising nitrogen fertiliser in landrace popcorn. *Rev. Ciência Agronômica*, 10.5935/1806-6690.20180032
- Massman, J.M., Jung, H.-J.G., Bernardo, R. (2013) Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. *Crop. Sci.* 53:58.
- Masuda, S., Mizusawa, K., Narisawa, T., Tozawa, Y., Ohta, H., Takamiya, K. (2008) The Bacterial Stringent Response, Conserved in Chloroplasts, Controls Plant Fertilization. *Plant Cell Physiol.* 49:135-141.
- Matika, O., Riggio, V., Moizan, M.A., Law, A.S., Wong, R.P., Archibald, A.L., Bishop, S.C. (2016) Genome-wide association reveals QTL for growth , bone and in vivo carcass traits as assessed by computed tomography in Scottish Blackface lambs. *Genetics Selection Evolution*, 48(1):11.
- Meng, F., Jia, H., Ling, N., Xue, Y., Liu, H., Wang, K., Yin, J., Li, Y. (2013) Cloning and characterization of two Argonaute genes in wheat (*Triticum aestivum* L.). *BMC Plant. Biol.* 13:18.
- Mengesha, W.A., Menkir, A., Unakchukwu, N., Meseka, S., Farinola, A., Girma, G., Gedil, M. (2017) Genetic diversity of tropical maize inbred lines combining resistance to *Striga hermonthica* with drought tolerance using SNP markers. *Plant. Breed.* 136:338-343.
- Mo, H., Wang, X., Zhang, Y., Zhang, G., Zhang, J., Ma, Z. (2015) Cotton polyamine oxidase is required for spermine and camalexin signalling in the defence response to *Verticillium dahliae*. *Plant J.* 83:962-975.
- Mora, F., Castillo, D., Lado, B., Matus, I., Poland, J., Belzile, F., von Zitzewitz, J., del Pozo, A. (2015) Genome-wide association mapping of agronomic traits

- and carbon isotope discrimination in a worldwide germplasm collection of spring wheat using SNP markers. *Mol. Breed.* 35:69.
- Morosini, J.S., Mendonça, L. de F., Lyra, D.H., Galli, G., Vidotti, M.S., Fritsch-Neto, R. (2017) Association mapping for traits related to nitrogen use efficiency in tropical maize lines under field conditions. *Plant Soil* 421:453-463.
- Mosher, R.A. (2010) Maternal control of Pol IV-dependent siRNAs in *Arabidopsis* endosperm. *New Phytol.* 186:358-364.
- Mosher, R.A., Melnyk, C.W., Kelly, K.A., Dunn, R.M., Studholme, D.J., Baulcombe, D.C. (2009) Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature*, 460:283-286.
- Mulamba, N.N., Mock, J.J. (1978) Improvement of yield potential of the ETO blanco maize (*Zea mays* L.) population by breeding for plant traits. *Egypt. J. Genet. Cytol.* 7:40-51.
- Mun, B.G., Lee, S.U., Park, E.J., Kim, H.H., Hussain, A., Imran, Q.M., Lee, I.J., Yun, B.W. (2017) Analysis of transcription factors among differentially expressed genes induced by drought stress in *Populus davidiana*. *3 Biotech*, 7(3):209.
- Mundim, G.B., Viana, J.M.S., Maia, C., Paes, G.P., DeLima, R.O., Valente, M.S.F. (2015) Inferring tropical popcorn gene pools based on molecular and phenotypic data. *Euphytica*, 202:55-68.
- Muñoz-Diez, C., Vitte, C., Ross-Ibarra, J., Gaut, B.S., Tenailon, M.I. (2012) Using Nextgen Sequencing to Investigate Genome Size Variation and Transposable Element Content. *Springer*, 24:41-58.
- Murphy, A.S., Hoogner, K.R., Peer, W.A., Taiz, L. (2002) Identification, Purification, and Molecular Cloning of N-1-Naphthylphthalamic Acid-Binding Plasma Membrane-Associated Aminopeptidases from *Arabidopsis*. *Plant Physiol*, 128:935-950.
- Nagamine, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Rudan, I.,

- Campbell, H., Wilson, J., Wild, S., Hicks, A.A., Pramstaller, P.P., Hastie, N., Wright, A.F., Haley, C.S. (2012) Localising Loci underlying Complex Trait Variation Using Regional Genomic Relationship Mapping. *PLoS One*, 7:e46501.
- Neves, L.G., Davis, J.M., Barbazuk, W.B., Kirst, M. (2014) A high-density gene map of loblolly pine (*Pinus taeda* L.) based on exome sequence capture genotyping. *G3: Genes, Genomes, Genetics*, 4:29-37.
- Newell, M.A., Cook, D., Tinker, N.A., Jannink, J.-L. (2011) Population structure and linkage disequilibrium in oat (*Avena sativa* L.): implications for genome-wide association studies. *Theor. Appl. Genet.* 122:623-632.
- Nziengui, H., Schoefs, B. (2009) Functions of reticulons in plants: What we can learn from animals and yeasts. *Cell Mol. Life Sci.* 66:584-595.
- Okeke, U.G., Akdemir, D., Rabbi, I., Kulakow, P., Jannink, J.-L. (2018) Regional Heritability Mapping Provides Insights into Dry Matter Content in African White and Yellow Cassava Populations. *The plant genome*, 11(1):1-18.
- Oliveira, N.C. de., Suzukawa, A.K., Pereira, C.B., Santos, H.V., Hanel, A., Albuquerque, F.A. de., Scapim, C.A. (2018) Popcorn genotypes resistance to fall armyworm. *Ciência Rural*, 10.1590/0103-8478cr20170378
- Ostertag, M., Stammler, J., Douchkov, D., Eichmann, R., Hückelhoven, R. (2013) The conserved oligomeric Golgi complex is involved in penetration resistance of barley to the barley powdery mildew fungus. *Mol. Plant. Pathol.* 14:230-240.
- Overby, C.L., Hripcsak, G., Shen, Y. (2015) Estimating heritability of drug-induced liver injury from common variants and implications for future study designs. *Sci. Rep.* 4:5762.
- Pace, J., Gardner, C., Romay, C., Ganapathysubramanian, B., Lübberstedt, T. (2015) Genome-wide association analysis of seedling root development in maize (*Zea mays* L.). *BMC Genomics*, 16:47.
- Pacheco, C.A.P., Gomes, E.E., Guimarães, P.E.O., Santos, M.X., Ferreira, A.S.

- (1998) Estimativas de parâmetros genéticos nas populações CMS-42 e CMS-43 de milho pipoca. *Pesqui. Agropecuária Bras.* 33:1995-2001.
- Paes, G.P., Viana, J.M.S., Silva, F.F. e., Mundim, G.B., Paes, G.P., Viana, J.M.S., Silva, F.F. e., Mundim, G.B. (2016) Linkage disequilibrium, SNP frequency change due to selection, and association mapping in popcorn chromosome regions containing QTLs for quality traits. *Genet. Mol. Biol.* 39:97-110.
- País, S.M., Téllez-Iñón, M.T., Capiati, D.A. (2009) Serine/Threonine Protein Phosphatases type 2A and their roles in stress signaling. *Plant Signal Behav.* 4:1013-1015.
- Pandey, M.K., Upadhyaya, H.D., Rathore, A., Vadez, V., Sheshshayee, M.S., Sriswathi, M., Govil, M., Kumar, A., Gowda, M.V.C., Sharma, S., Hamidou, F., Kumar, V.A., Khera, P., Bhat, R.S., Khan, A.W., Singh, S., Li, H., Monyo, E., Nadaf, H.L., Mukri, G., Jackson, S.A., Guo, B., Liang, X., Varshney, R.K. (2014) Genomewide Association Studies for 50 Agronomic Traits in Peanut Using the 'Reference Set' Comprising 300 Genotypes from 48 Countries of the Semi-Arid Tropics of the World. *PLoS One*, 9:e105228.
- Pereira, M.G., Amaral Júnior, A.T. (2001) Estimation of Genetic Components in Popcorn Based on the Nested Design. *Crop. Breed. Appl. Biotechnol.* 1:3-10.
- Perera, I.Y., Hung, C.-Y., Moore, C.D., Stevenson-Paulik, J., Boss, W.F. (2008) Transgenic *Arabidopsis* Plants Expressing the Type 1 Inositol 5-Phosphatase Exhibit Increased Drought Tolerance and Altered Abscisic Acid Signaling. *Plant Cell*, 20:2876-2893.
- Perez, D.E., Hoyer, J.S., Johnson, A.I., Moody, Z.R., Lopez, J., Kaplinsky, N.J. (2009) BOBBER1 Is a Noncanonical *Arabidopsis* Small Heat Shock Protein Required for Both Development and Thermotolerance. *Plant Physiol.* 151:241-252.
- Pootakham, W., Jomchai, N., Ruang-areerate, P., Shearman, J.R., Sonthirod, C., Sangsrakru, D., Tragoonrung, S., Tangphatsornruang, S. (2015) Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). *Genomics*, 105:288-295.

- Portwood, J.L., Woodhouse, M.R., Cannon, E.K., Gardiner, J.M., Harper, L.C., Schaeffer, M.L., Walsh, J.R., Sen, T.Z., Cho, K.T., Schott, D.A., Braun, B.L., Dietze, M., Dunfee, B., Elsik, C.G., Manchanda, N., Coe, E., Sachs, M., Stinard, P., Tolbert, J., Zimmerman, S., Andorf, C.M. (2019) MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic. Acids Res.* 47:D1146-D1154.
- Price, A.L., Zaitlen, N.A., Reich, D., Patterson, N. (2010) New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* 11:459-463.
- Pritchard, J.K., Stephens, M., Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155:945-59.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C. (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81:559-575.
- Qi, Y., Liu, Y., Zhang, Z., Gao, J., Guan, Z., Fang, W., Chen, S., Chen, F., Jiang, J. (2018) The over-expression of a chrysanthemum gene encoding an RNA polymerase II CTD phosphatase-like 1 enzyme enhances tolerance to heat stress. *Hortic. Res.* 5:37.
- Qin, N., Xu, W., Hu, L., Li, Y., Wang, H., Qi, X., Fang, Y., Hua, X. (2016) Drought tolerance and proteomics studies of transgenic wheat containing the maize C4 phosphoenolpyruvate carboxylase (PEPC) gene. *Protoplasma*, 253:1503-1512.
- R Core Team. (2016) R: A Language and Environment for Statistical Computing.
- Rahim, F., Almeida, V.C., Viana, J.M.S., Ribeiro, C., Risso, L.A., Ribeiro, M.P. (2019) Identification of contrasting tropical popcorn inbreds for studying aluminum toxicity tolerance inheritance. *Euphytica*, 215:47.
- Rangel, R.M., Amaral Júnior, A. do., Gonçalves, L.S.A., Freitas Júnior, S. de P., Candido, L.S. (2011) Análise biométrica de ganhos por seleção em população de milho pipoca de quinto ciclo de seleção recorrente. *Rev.*

Ciência Agronômica, 42:473-481.

- Registro Nacional de Cultivares. (2019) Registro Nacional de Cultivares - RNC: <http://www.agricultura.gov.br/guia-de-servicos/registro-nacional-de-cultivares-rnc> em 17/10/19.
- Resende, R.T. (2017) *Regional heritability mapping and gwas for molecular breeding in eucalyptus hybrids*. Tese (Doutorado em Genética e Melhoramento) – Viçosa – MG, Universidade Federal de Viçosa –UFV, 59p.
- Resende, R.T., de Resende, M.D. V., Azevedo, C.F., Fonseca e Silva, F., Melo, L.C., Pereira, H.S., Souza, T.L.P.O., Valdisser, P.A.M.R., Brondani, C., Vianello, R.P. (2018) Genome-wide association and regional heritability mapping of plant architecture, lodging and productivity in *Phaseolus vulgaris*. *G3: Genes, Genomes, Genetics*, 8(8):2841-2854.
- Resende, R.T., Resende, M.D.V., Silva, F.F., Azevedo, C.F., Takahashi, E.K., Silva-Junior, O.B., Grattapaglia, D. (2017) Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in Eucalyptus. *New Phytol.* 213:1287-1300.
- Revilla, P., Rodríguez, V.M., Ordás, A., Rincent, R., Charcosset, A., Giauffret, C., Melchinger, A.E., Schön, C.-C., Bauer, E., Altmann, T., Brunel, D., Moreno-González, J., Campo, L., Ouzunova, M., Álvarez, Á., Ruíz de Galarreta, J.I., Laborde, J., Malvar, R.A. (2016) Association mapping for cold tolerance in two large maize inbred panels. *BMC Plant. Biol.* 16:127.
- Ribeiro, R.M., Amaral Júnior, A.T. do., Pena, G.F., Vivas, M., Kurosawa, R.N., Gonçalves, L.S.A. (2016) Effect of recurrent selection on the variability of the UENF-14 popcorn population. *Crop. Breed. Appl. Biotechnol.* 16:123-131.
- Ribeiro, R.M., do Amaral Júnior, A.T., Gonçalves, L.S.A., Candido, L.S., Silva, T.R.C., Pena, G.F. (2012) Genetic progress in the UNB-2U population of popcorn under recurrent selection in Rio de Janeiro, Brazil. *Genet. Mol. Res.* 11:1417-1423.
- Rigas, S., Daras, G., Tsitsekian, D., Alatzas, A., Hatzopoulos, P. (2014) Evolution and significance of the Lon gene family in *Arabidopsis* organelle biogenesis

- and energy metabolism. *Front. Plant. Sci.*10.3389/fpls.2014.00145
- Riggio, V., Matika, O., Pong-Wong, R., Stear, M.J., Bishop, S.C. (2013) Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. *Heredity (Edinb)*, 110:420-429.
- Rocha, J.A., Vasconcelos, S., da Silva, F.M.M., Jurkiewicz Melo, A., Souza Silva, M.F., de Miranda, J.A.L., Benko-Iseppon, A.M., de Andrade, I.M. (2014) ISSR Primer Selection for Genetic Variability Analyses with Jaborandi (*Pilocarpus microphyllus* Stapf ex Wardlew., Rutaceae). *For. Res.: Open. Access*, 3(4).
- Rodvalho, M. de A., Mora, F., Santos, E.M. dos., Scapim, C.A., Arnhold, E. (2008) Heredabilidad de la sobrevivencia en 169 familias de maíz-roseta de granos blancos: Un enfoque Bayesiano. *Cienc Investig. Agrar.* 35:303-309.
- Rossi, E.S., Kuki, M.C., Pinto, R.J.B., Scapim, C.A., Faria, M. V., Leon, N. (2020) Genomic-wide association study for white spot resistance in a tropical maize germplasm. *Euphytica*, 216:15.
- Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K., Yamaguchi-Shinozaki, K. (2002) DNA-Binding Specificity of the ERF/AP2 Domain of *Arabidopsis* DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. *Biochem. Biophys. Res. Commun.* 290:998-1009.
- Sanches, R.E., Suzukawa, A.K., Contreras-Soto, R.I., Rizzardi, D.A., Kuki, M.C., Zeffa, D.M., Albuquerque, F.A. de., Scapim, C.A. (2019) Multivariate analysis reveals key traits of fall armyworm resistance in tropical popcorn genotypes. *Bragantia*, 78:175-182.
- Santos, A., Amaral Júnior, A.T., Fritsche-Neto, R., Kamphorst, S.H., Ferreira, F.R.A., Amaral, J.F.T., Vivas, J.M.S., Santos, P.H.A.D., Lima, V.J., Khan, S., Schmitt, K.F.M., Leite, J.T., Junior, D.R.S., Bispo, R.B., Santos, T.O., Oliveira, U.A., Guimarães, L.J.M., Rodriguez, O. (2019) Relative importance of gene effects for nitrogen-use efficiency in popcorn. *PLoS One*, 14:e0222726.
- Santos, A., Amaral Júnior, A.T., Kurosawa, R.N.F., Gerhardt, I.F.S., Fritsche Neto, R. (2017a) GGE Biplot projection in discriminating the efficiency of popcorn

- lines to use nitrogen. *Ciência e Agrotecnologia*, 41: 22–31.
- Santos, F.S., Amaral Júnior, A.T. do., Freitas Júnior, S. de P., Rangel, R.M., Pereira, M.G. (2007) Predição de ganhos genéticos por índices de seleção na população de milho pipoca UNB-2U sob seleção recorrente. *Bragantia*, 66:389-396.
- Santos, J.S., Amaral Júnior, A.T., Vivas, M., Mafra, G.S., Pena, G.F., Silva, F.H.L., Guimarães, A.G. (2017b) Genetic control and combining ability of agronomic attributes and northern leaf blight-related attributes in popcorn. *Genet. Mol. Res.* 16(3).
- Sawazaki, E. (2010) Milho pipoca. *In: An. do XXVIII Congr. Nac. Milho e Sorgo Milho e Sorgo*.
- Sawazaki, E. (1996) Parâmetros genéticos em milho pipoca (*Zea mays* L.). Tese (Doutorado em Agronomia) – Piracicaba – SP, Escola Superior de Agricultura Luiz de Queiroz/USP – ESALQ/USP, 157p.
- Sawazaki, E., Castro, J.L. de., Sordi, G. de., Fantin, G.M., Dudienas, C., Lordello, A.I.L., Lordello, R.R.A. (1998) Características agronômicas e de fitossanidade de híbridos exóticos de milho pipoca. *In: Congresso Nacional de Milho e Sorgo*.
- Sawazaki, E., Gallo, P.B., Sordi, G.D., Longo, L.S. (1986) Estudo da capacidade de expansão em cruzamentos dialélicos entre variedades de milho-pipoca. *In: Anais do XV Congresso Nacional de Milho e Sorgo*.
- Sawazaki, E., Paterniani, M. elisa A.G.Z., Castro, J.L. de., Gallo, P.B., Galvão, J.C.C., SaeS, L.A. (2000) Potencial de linhagens de populações locais de milho pipoca para síntese de híbridos. *Bragantia*, 59:143-151.
- Scapim, C.A., Pacheco, C.A.P., Tonet, A., Braccini, A.D.L., Pinto, R.J.B. (2002) Análise dialélica e heterose de populações de milho-pipoca. *Bragantia*, 61:219-230.
- Schäfer-Pregl, R., Ritter, E., Concilio, L., Hesselbach, J., Lovatti, L., Walkemeier, B., Thelen, H., Salamini, F., Gebhardt, C. (1998) Analysis of quantitative trait

- loci (QTLs) and quantitative trait alleles (QTAs) for potato tuber yield and starch content. *Theor. Appl. Genet.* 97:834-846.
- Schmitt, K.F.M., Lima, V.J., Júnior, A.T.A., Santos, J.S., Mafra, G.S., Vivas, M., Kamphorst, S.H., Souza, Y.P., Oliveira, F.T., Ferreira, F.R.A., Vivas, J.M.S. (2019) Research Article Combining ability of popcorn lines for resistance to the fungus *Puccinia polysora* (Pucciniaceae). *Genet. Mol. Res.* 18(3):gmr18330
- Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Wilson, R.K., et al. (2009) The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science*, 326(5956):1112-1115.
- Schwantes, I.A., Amaral Júnior, A.T., Vivas, M., Almeida Filho, J.E., Kamphorst, S.H., Guimarães, A.G., Khan, S. (2018) Inheritance of resistance to Fusarium ear rot in popcorn. *Crop. Breed. Appl. Biotechnol.* 18:81-88.
- Senhorinho, H.J.C., Coan, M.M.D., Marino, T.P., Kuki, M.C., Pinto, R.J.B., Scapim, C.A., Holland, J.B. (2019) Genomic-Wide Association Study of Popping Expansion in Tropical Popcorn and Field Corn Germplasm. *Crop. Sci.* 59:2007.
- Serba, D.D., Daverdin, G., Bouton, J.H., Devos, K.M., Brummer, E.C., Saha, M.C. (2015) Quantitative Trait Loci (QTL) Underlying Biomass Yield and Plant Height in Switchgrass. *BioEnergy Res.* 8:307-324.
- Sharma, M., Pandey, G.K. (2016) Expansion and Function of Repeat Domain Proteins During Stress and Development in Plants. *Front. Plant. Sci.* 10.3389/fpls.2015.01218
- Sharma, N., Park, S.-W., Vepachedu, R., Barbieri, L., Ciani, M., Stirpe, F., Savary, B.J., Vivanco, J.M. (2004) Isolation and Characterization of an RIP (Ribosome-Inactivating Protein)-Like Protein from *Tobacco* with Dual Enzymatic Activity. *Plant. Physiol.* 134:171-181.
- Sharoni, A.M., Nuruzzaman, M., Satoh, K., Shimizu, T., Kondoh, H., Sasaya, T., Choi, I.-R., Omura, T., Kikuchi, S. (2011) Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice.

Plant Cell Physiol. 52:344-360.

- She, K.C., Kusano, H., Koizumi, K., Yamakawa, H., Hakata, M., Imamura, T., Fukuda, M., Naito, N., Tsurumaki, Y., Yaeshima, M., Tsuge, T., Matsumoto, K., Kudoh, M., Itoh, E., Kikuchi, S., Kishimoto, N., Yazaki, J., Ando, T., Yano, M., Aoyama, T., Sasaki, T., Satoh, H., Shimada, H. (2010) A Novel Factor FLOURY ENDOSPERM2 Is Involved in Regulation of Rice Grain Size and Starch Quality. *Plant Cell*, 22:3280-3294.
- Shirali, M., Pong-Wong, R., Navarro, P., Knott, S., Hayward, C., Vitart, V., Rudan, I., Campbell, H., Hastie, N.D., Wright, A.F., Haley, C.S. (2016) Regional heritability mapping method helps explain missing heritability of blood lipid traits in isolated populations. *Heredity (Edinb)*, 116:333-338.
- Silva, T.R. da C., Teixeira do Amaral Júnior, A., Eustáquio de Almeida Filho, J., Simone Mendonça Freitas, M., Gonçalves Guimarães, A., Henrique Kamphorst, S. (2019) Contrasting phosphorus environments as indicators for popcorn breeding lines. *Funct. Plant .Breed. J.* 1:1-15.
- Simko, I., Haynes, K.G., Jones, R.W. (2006) Assessment of Linkage Disequilibrium in Potato Genome With Single Nucleotide Polymorphism Markers. *Genetics*, 173:2237-2245.
- Singh, S.K., Ram, U.S., Singh, M.K., Deshmukh, R. (2017) Effect of Planting Time, Fertility Level and Plant Population on Development, Yield, Nutrient Uptake and Quality of Winter Popcorn (*Zea mays everta* Sturt) under Late Sown Condition. *Int. J. Curr. Microbiol. Appl. Sci.* 6:1187-1193.
- Smith, H.F. (1936) A discriminant function for plant selection. *Ann. Eugen.* 7:240-250.
- Soni, N. V., Khanorkar, S.M. (2014) Genetic architecture of yield traits and popping quality in popcorn (*Zea mays* var. *everta*) inbred lines. *Electron. J. Plant. Breed.* 5:1-16.
- Sonoda, Y., Ikeda, A., Saiki, S., Wirén, N. von., Yamaya, T., Yamaguchi, J. (2003) Distinct Expression and Function of Three Ammonium Transporter Genes (OsAMT1;1 – 1;3) in Rice. *Plant Cell Physiol.* 44:726-734.

- Springer, N.M., Anderson, S.N., Andorf, C.M., Ahern, K.R., Bai, F., Barad, O., Brutnell, T.P., et al. (2018) The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat. Genet.* 50:1282-1288.
- Stirpe, F. (2013) Ribosome-inactivating proteins: From toxins to useful proteins. *Toxicon*, 67:12-16.
- Su, J., Fan, S., Li, L., Wei, H., Wang, C., Wang, H., Song, M., Zhang, C., Gu, L., Zhao, S., Mao, G., Wang, C., Pang, C., Yu, S. (2016) Detection of Favorable QTL Alleles and Candidate Genes for Lint Percentage by GWAS in Chinese Upland Cotton. *Front. Plant. Sci.* 10.3389/fpls.2016.01576
- Sun, R., Ye, R., Gao, L., Zhang, L., Wang, R., Mao, T., Zheng, Y., Li, D., Lin, Y. (2017) Characterization and Ectopic Expression of CoWRI1, an AP2/EREBP Domain-Containing Transcription Factor from Coconut (*Cocos nucifera* L.) Endosperm, Changes the Seeds Oil Content in Transgenic *Arabidopsis thaliana* and Rice (*Oryza sativa* L.). *Front Plant Sci.* 8:63.
- Sun, S., Zhou, Y., Chen, J., Shi, J., Zhao, H., Zhao, H., Song, W., Zhang, M., Cui, Y., Dong, X., Liu, H., Ma, X., Jiao, Y., Wang, B., Wei, X., Stein, J.C., Glaubitz, J.C., Lu, F., Yu, G., Liang, C., Fengler, K., Li, B., Rafalski, A., Schnable, P.S., Ware, D.H., Buckler, E.S., Lai, J. (2018) Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. *Nat. Genet.* 50:1289-1295.
- Tai, F.J., Yuan, Z.L., Wu, X.L., Zhao, P.F., Hu, X.L., Wang, W. (2011) Identification of membrane proteins in maize leaves, altered in expression under drought stress through polyethylene glycol treatment. *Plant. Omi. J.* 4:250-256.
- Tenaillon, M.I., Manicacci, D., Nicolas, S.D., Tardieu, F., Welcker, C. (2016) Testing the link between genome size and growth rate in maize. *PeerJ*, 4:e2408.
- Terra, T.G.R., Leal, T.C.A. de B., Borém, A., Rangel, P.H.N. (2013) Tolerância de linhagens de arroz de terras altas à seca. *Pesqui Agropecuária Trop* 43: 201–208.
- Thornton, K.R., Foran, A.J., Long, A.D. (2013) Properties and Modeling of GWAS

- when Complex Disease Risk Is Due to Non-Complementing, Deleterious Mutations in Genes of Large Effect. *PLoS Genet.* 9:e1003258.
- Thorwarth, P., Yousef, E.A.A., Schmid, K.J. (2018) Genomic Prediction and Association Mapping of Curd-Related Traits in Gene Bank Accessions of Cauliflower. *G3: Genes, Genomes, Genet.* 8:707-718.
- Tian, S., Muneeruddin, K., Choi, M.Y., Tao, L., Bhuiyan, R.H., Ohmi, Y., Furukawa, K., Furukawa, K., Boland, S., Shaffer, S.A., Adam, R.M., Dong, M. (2018) Genome-wide CRISPR screens for Shiga toxins and ricin reveal Golgi proteins critical for glycosylation. *PLOS Biol.* 16:e2006951.
- Toledo, E.R. de., Leandro, R.A., Souza Junior, C.L. de., Souza, A.P. de. (2008) Mapeamento de QTLs: uma abordagem bayesiana. *Rev. Bras. Biom.* 26:107-114.
- Tolley, N., Sparkes, I.A., Hunter, P.R., Craddock, C.P., Nuttall, J., Roberts, L.M., Hawes, C., Pedrazzini, E., Frigerio, L. (2008) Overexpression of a Plant Reticulon Remodels the Lumen of the Cortical Endoplasmic Reticulum but Does not Perturb Protein Transport. *Traffic*, 9:94-102.
- Uemoto, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Wilson, J.F., Rudan, I., Campbell, H., Hastie, N.D., Wright, A.F., Haley, C.S. (2013) The power of regional heritability analysis for rare and common variant detection: simulations and application to eye biometrical traits. *Front. Genet.* 4:232.
- Van Inghelandt, D., Melchinger, A.E., Lebreton, C., Stich, B. (2010) Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. *Theor. Appl. Genet.* 120:1289-1299.
- VanRaden, P.M. (2008) Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* 91:4414-4423.
- Vaucheret, H. (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.* 18:1187-1197.
- Vaucheret, H., Mallory, A.C., Bartel, D.P. (2006) AGO1 Homeostasis Entails

- Coexpression of MIR168 and AGO1 and Preferential Stabilization of miR168 by AGO1. *Mol. Cell*, 22:129-136.
- Venancio, T.M., Aravind, L. (2010) CYSTM, a novel cysteine-rich transmembrane module with a role in stress tolerance across eukaryotes. *Bioinformatics*, 26:149-152.
- Viana, J.M.S., Silva, F.F., Mundim, G.B., Azevedo, C.F., Jan, H.U. (2017) Efficiency of low heritability QTL mapping under high SNP density. *Euphytica*, 213:13.
- Viana, J.M.S., Mundim, G.B., Lima, R.O., Silva, F.F., Resende, M.D. V. (2014) Best linear unbiased prediction for genetic evaluation in reciprocal recurrent selection with popcorn populations. *J. Agric. Sci.* 152:428-438.
- Vittorazzi, C., Amaral Júnior, A.T., Candido, L.S., Freitas, I.L. de J., Silva, T.R. da C. (2017) Arranjo populacional para a variedade UENF-14 de milho pipoca. *Rev. Bras. Milho e Sorgo*, 16:401.
- Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., Dehesh, K. (2010) Arabidopsis Deadenylation AtCAF1a and AtCAF1b Play Overlapping and Distinct Roles in Mediating Environmental Stress Responses. *Plant Physiol.* 152:866-875.
- Wang, H., Misztal, I., Aguilar, I., Legarra, A., Fernando, R.L., Vitezica, Z., Okimoto, R., Wing, T., Hawken, R., Muir, W.M. (2014) Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. *Front. Genet.* 94(2):73-83.
- Wang, L., Song, X., Gu, L., Li, X., Cao, S., Chu, C., Cui, X., Chen, X., Cao, X. (2013) NOT2 Proteins Promote Polymerase II-Dependent Transcription and Interact with Multiple MicroRNA Biogenesis Factors in *Arabidopsis*. *Plant Cell* 25:715-727.
- Wang, M., Yan, J., Zhao, J., Song, W., Zhang, X., Xiao, Y., Zheng, Y. (2012) Genome-wide association study (GWAS) of resistance to head smut in maize. *Plant. Sci.* 196:125-131.

- Wang, S., Feng, J., Ren, W., Huang, B., Zhou, L., Wen, Y.-J., Zhang, J., Dunwell, J.M., Xu, S., Zhang, Y. (2016) Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* 6:19444.
- Wang, X. (2009) Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS Lett.* 583:3303-3309.
- Wayne, L.L., Wallis, J.G., Kumar, R., Markham, J.E., Browse, J. (2013) Cytochrome b5 Reductase Encoded by CBR1 Is Essential for a Functional Male Gametophyte in Arabidopsis. *Plant Cell* 25:3052-3066.
- Wen, W., Li, D., Li, X., Gao, Y., Li, W., Li, H., Liu, J., Liu, H., Chen, W., Luo, J., Yan, J. (2014) Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun* 5:3438.
- Wen, Y.-J., Zhang, H., Ni, Y.-L., Huang, B., Zhang, J., Feng, J.-Y., Wang, S.-B., Dunwell, J.M., Zhang, Y.-M., Wu, R. (2018) Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Brief Bioinform.* 19:700-712.
- Weng, J., Xie, C., Hao, Z., Wang, J., Liu, C., Li, M., Zhang, D., Bai, L., Zhang, S., Li, X. (2011) Genome-Wide Association Study Identifies Candidate Genes That Affect Plant Height in Chinese Elite Maize (*Zea mays* L.) Inbred Lines. *PLoS One*, 6:e29229.
- Wu, M.C., Lee, S., Cai, T., Li, Y., Boehnke, M., Lin, X. (2011) Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. *Am. J. Hum. Genet.* 89:82-93.
- Wu, Y., Pu, C., Lin, H., Huang, H., Huang, Y., Hong, C., Chang, M., Lin, Y. (2015) Three novel alleles of FLOURY ENDOSPERM2 (FLO2) confer dull grains with low amylose content in rice. *Plant. Sci.* 233:44-52.
- Xiao, Y., Liu, H., Wu, L., Warburton, M., Yan, J. (2017) Genome-wide Association Studies in Maize: Praise and Stargaze. *Mol. Plant.* 10:359-374.
- Xu, C., Ren, Y., Jian, Y., Guo, Z., Zhang, Y., Xie, C., Fu, J., Wang, H., Wang, G.,

- Xu, Y., Li, P., Zou, C. (2017) Development of a maize 55 K SNP array with improved genome coverage for molecular breeding. *Mol. Breed.* 37:20.
- Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E., Visscher, P.M. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42:565-569.
- Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M. (2011) GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* 88:76-82.
- Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M., Price, A.L. (2014a) Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46:100-106.
- Yang, N., Liu, J., Gao, Q., Gui, S., Chen, L., Yang, L., Huang, J., Deng, T., Luo, J., He, L., Wang, Y., Xu, P., Peng, Y., Shi, Z., Lan, L., Ma, Z., Yang, X., Zhang, Q., Bai, M., Li, S., Li, W., Liu, L., Jackson, D., Yan, J. (2019) Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. *Nat. Genet.* 51:1052-1059.
- Yang, N., Lu, Y., Yang, X., Huang, J., Zhou, Y., Ali, F., Wen, W., Liu, J., Li, J., Yan, J. (2014b) Genome Wide Association Studies Using a New Nonparametric Model Reveal the Genetic Architecture of 17 Agronomic Traits in an Enlarged Maize Association Panel. *PLoS Genet.* 10:e1004573.
- Yang, N., Xu, X., Wang, R., Peng, W., Cai, L., Song, J.-M., Li, W., Luo, X., Niu, L., Wang, Y., Jin, M., Chen, L., Luo, J., Deng, M., Wang, L., Pan, Q., Liu, F., Jackson, D., Yang, X., Chen, L., Yan, J. (2017) Contributions of *Zea mays* subspecies mexicana haplotypes to modern maize. *Nat. Commun* 8:1874.
- Yoshioka, H., Numata, N., Nakajima, K., Katou, S., Kawakita, K., Rowland, O., Jones, J.D.G., Doke, N. (2003) Nicotiana benthamiana gp91 phox Homologs NbrbohA and NbrbohB Participate in H₂O₂ Accumulation and Resistance to Phytophthora infestans. *Plant Cell.* 15:706-718.
- Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S.,

- Buckler, E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38:203-208.
- Yu, R.M.K., Zhou, Y., Xu, Z.-F., Chye, M.-L., Kong, R.Y.C. (2003) Two genes encoding protein phosphatase 2A catalytic subunits are differentially expressed in rice. *Plant. Mol. Biol.* 51:295–311.
- Yu, Y., Zhao, Z., Shi, Y., Tian, H., Liu, L., Bian, X., Xu, Y., Zheng, X., Gan, L., Shen, Y., Wang, C., Yu, X., Wang, C., Zhang, X., Guo, X., Wang, J., Ikehashi, H., Jiang, L., Wan, J. (2016) Hybrid Sterility in Rice (*Oryza sativa* L.) Involves the Tetratricopeptide Repeat Domain Containing Protein. *Genetics*, 203:1439-1451.
- Zain, N.A.M., Ismail, M.R., Puteh, A., Mahmood, M., Islam, M.R. (2014) Impact of cyclic water stress on growth, physiological responses and yield of rice (*Oryza sativa* L.) grown in tropical environment. *Ciência Rural*, 44:2136-2141.
- Zaitlen, N., Kraft, P. (2012) Heritability in the genome-wide association era. *Hum Genet.* 131:1655-1664.
- Zeng, A., Chen, P., Korth, K., Hancock, F., Pereira, A., Brye, K., Wu, C., Shi, A. (2017a) Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. *Mol. Breed.* 37:30.
- Zeng, Y., Navarro, P., Shirali, M., Howard, D.M., Adams, M.J., Hall, L.S., Clarke, T.-K., Thomson, P.A., Smith, B.H., Murray, A., Padmanabhan, S., Hayward, C., Boutin, T., MacIntyre, D.J., Lewis, C.M., Wray, N.R., Mehta, D., Penninx, B.W.J.H., Milaneschi, Y., Baune, B.T., Air, T., Hottenga, J.-J., Mbarek, H., Castelao, E., Pistis, G., Schulze, T.G., Streit, F., Forstner, A.J., Byrne, E.M., Martin, N.G., Breen, G., Müller-Myhsok, B., Lucae, S., Kloiber, S., Domenici, E., Deary, I.J., Porteous, D.J., Haley, C.S., McIntosh, A.M. (2017b) Genome-wide Regional Heritability Mapping Identifies a Locus Within the TOX2 Gene Associated With Major Depressive Disorder. *Biol. Psychiatry*, 82:312-321.
- Zeng, Z.B. (1994) Precision mapping of quantitative trait loci. *Genetics*, 136:1457-68.
- Zhai, L., Sun, W., Zhang, K., Jia, H., Liu, L., Liu, Z., Teng, F., Zhang, Z. (2014)

- Identification and characterization of Argonaute gene family and meiosis-enriched Argonaute during sporogenesis in maize. *J. Integr. Plant. Biol.* 56:1042-1052.
- Zhai, L., Wang, L., Teng, F., Zhou, L., Zhang, W., Xiao, J., Liu, Y., Deng, W. (2016) Argonaute and Argonaute-Bound Small RNAs in Stem Cells. *Int. J. Mol. Sci.* 17:208.
- Zhang, D., Jiang, P., Xu, Q., Zhang, X. (2011) Arginine and Glutamate-rich 1 (ARGLU1) Interacts with Mediator Subunit 1 (MED1) and Is Required for Estrogen Receptor-mediated Gene Transcription and Breast Cancer Cell Growth. *J. Biol. Chem.* 286:17746-17754.
- Zhang, J., Song, Q., Cregan, P.B., Nelson, R.L., Wang, X., Wu, J., Jiang, G.-L. (2015a) Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics*, 16:217.
- Zhang, L., Tian, L.-H., Zhao, J.-F., Song, Y., Zhang, C.-J., Guo, Y. (2009) Identification of an Apoplastic Protein Involved in the Initial Phase of Salt Stress Response in Rice Root by Two-Dimensional Electrophoresis. *Plant Physiol.* 149:916-928.
- Zhang, M., Chen, Y., Jin, X., Cai, Y., Yuan, Y., Fu, C., Yu, L. (2019) New different origins and evolutionary processes of AP2/EREBP transcription factors in *Taxus chinensis*. *BMC Plant Biol.* 19:413.
- Zhang, M., Wang, C., Lin, Q., Liu, A., Wang, T., Feng, X., Liu, J., Han, H., Ma, Y., Bonea, D., Zhao, R., Hua, X. (2015b) A tetratricopeptide repeat domain-containing protein SSR1 located in mitochondria is involved in root development and auxin polar transport in *Arabidopsis*. *Plant J.* 83:582-599.
- Zhang, X., Zhang, H., Li, L., Lan, H., Ren, Z., Liu, D., Wu, L., Liu, H., Jaqueth, J., Li, B., Pan, G., Gao, S. (2016a) Characterizing the population structure and genetic diversity of maize breeding germplasm in Southwest China using genome-wide SNP markers. *BMC Genomics.* 17:697.
- Zhang, X., Zhang, J., Zhang, W., Yang, T., Xiong, Y., Che, D. (2016b)

- Transcriptome sequencing and de novo analysis of *Rosa multiflora* under cold stress. *Acta Physiol. Plant.* 38:164.
- Zhang, Y., Yu, C., Lin, J., Liu, J., Liu, B., Wang, J., Huang, A., Li, H., Zhao, T. (2017) OsMPH1 regulates plant height and improves grain yield in rice. *PLoS One*, 12:e0180825.
- Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., Bradbury, P.J., Yu, J., Arnett, D.K., Ordovas, J.M., Buckler, E.S. (2010) Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42:355-360.
- Zheng, K., Ni, Z., Qu, Y., Cai, Y., Yang, Z., Sun, G., Chen, Q. (2018) Genome-wide identification and expression analyses of TCP transcription factor genes in *Gossypium barbadense*. *Sci. Rep.* 8:14526.
- Zhong, J., He, W., Peng, Z., Zhang, H., Li, F., Yao, J. (2019) A putative AGO protein, OsAGO17, positively regulates grain size and grain weight through OsmiR397b in rice. *Plant Biotechnol. J.* 18(4):916-928.
- Zhou, S., Wei, F., Nguyen, J., Bechner, M., Potamouisis, K., Goldstein, S., Pape, L., Mehan, M.R., Churas, C., Pasternak, S., Forrest, D.K., Wise, R., Ware, D., Wing, R.A., Waterman, M.S., Livny, M., Schwartz, D.C. (2009) A Single Molecule Scaffold for the Maize Genome. *PLoS Genet.* 5:e1000711.
- Zhu, X.-M., Shao, X.-Y., Pei, Y.-H., Guo, X.-M., Li, J., Song, X.-Y., Zhao, M.-A. (2018) Genetic Diversity and Genome-Wide Association Study of Major Ear Quantitative Traits Using High-Density SNPs in Maize. *Front Plant. Sci.* 9:966.
- Zila, C.T., Samayoa, L.F., Santiago, R., Butrón, A., Holland, J.B. (2013) A Genome-Wide Association Study Reveals Genes Associated with Fusarium Ear Rot Resistance in a Maize Core Diversity Panel. *G3: Genes, Genomes, Genetics*, 3:2095-2104.
- Zuo, X., Jia, J. (2009) Promoter polymorphisms which modulate insulin degrading enzyme expression may increase susceptibility to Alzheimer's disease. *Brain. Res.* 1249:1-8.