Universidade de São Paulo
Escola Superior de Agricultura Luiz de Queiroz

**Mapeamento associativo para múltiplos ambientes e múltiplos locos,
visando tolerância à seca em milho**

Carina de Oliveira Anoni

Tese apresentada para obtenção do título de Doutora em
Ciências. Área de concentração: Genética e Melhora-
mento de Plantas

Piracicaba
2016

Carina de Oliveira Anoni
Engenheira Agrônoma

**Mapeamento associativo para múltiplos ambientes e múltiplos locos,
visando tolerância à seca em milho**

Orientador:

Prof. Dr. **ANTONIO AUGUSTO FRANCO GARCIA**

Tese apresentada para obtenção do título de Doutora em
Ciências. Área de concentração: Genética e Melhora-
mento de Plantas

Piracicaba
2016

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Anoni, Carina de Oliveira

Mapeamento associativo para múltiplos ambientes e múltiplos locos, visando tolerância à seca em milho / Carina de Oliveira Anoni. -- Piracicaba, 2016 .
64 p.

Tese (Doutorado) -- USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. *Zea mays* 2. GWAS 3. *G×E* 4. Tolerância à seca . I. Título.

DEDICATÓRIA

*À minha avó **Maria**,
por seu amor,
ofereço.*

*À minha mãe, **Evaneide**, e ao meu pai **Luiz**,
por serem meus exemplos de vida.
Pelo amor incondicional e pelo apoio
em todos os sonhos e decisões de minha vida,
dedico.*

AGRADECIMENTOS

A Deus.

Ao Prof. Antonio Augusto Franco Garcia, que ao longo desses anos dedicou seu tempo em ensinar não apenas a ciência e pesquisa, mas também seus princípios de vida valiosos. Por depositar toda sua confiança em minha pessoa e acreditar em meu potencial. Serei eternamente grata pela oportunidade e receptividade desde o início dessa caminhada. Sua orientação e amizade contribuíram muito para minha formação profissional e pessoal.

Aos meus pais, pelo amor maior. Por muitas vezes abdicarem de seus próprios sonhos para que eu pudesse chegar até aqui. Por ensinarem desde sempre que a educação, juntamente com os nossos princípios são os bens mais valiosos que carregamos. Sem vocês nada disso faria sentido.

Ao meu irmão Ricardo, por sempre me escutar e acompanhar essa trajetória. Seus conselhos, seus abraços e até mesmo seu silêncio me fazem refletir sobre nossos valores.

Ao Renato, por seu amor e companheirismo. Seu apoio, carinho e paciência ímpar foram fundamentais para que eu conseguisse realizar esse sonho.

À ESALQ/USP e ao Departamento de Genética, pela oportunidade concedida.

À CAPES, pelas bolsas de doutorado no início do curso e sanduíche (010141/2014-06) e ao CNPq, pela bolsa de doutorado ao longo do curso (145478/2012-2).

À Embrapa Milho e Sorgo, Dra. Claudia Teixeira Guimarães, Dr. Lauro José Guimarães e suas respectivas equipes, pela oportunidade de colaboração e disponibilidade dos dados que foram essenciais para o desenvolvimento desse trabalho. Ao Kaio Olímpio Dias, pela troca de conhecimentos construtiva que tivemos ao longo dessa parceria. Em especial à Dra. Maria Marta Pastina, pela motivação de desenvolvimento desse trabalho, pelas conversas prazerosas e por compartilhar com grande entusiasmo suas ideias e experiências.

À Wageningen University, Biometris, pelo suporte durante o doutorado sanduíche. Em especial, ao Dr. Fred van Eeuwijk, Dr. Martin Boer e Dr. Marcos Malosetti, que compartilharam seus conhecimentos com muita dedicação e, além de se empenharem nesse projeto, contribuíram para a minha formação.

Aos colegas do Laboratório de Genética Estatística da ESALQ, à velha guarda: Maria Marta, Rodrigo, Graciela, Gabriel, Marcelo, Renato, Edjane; à jovem guarda: João Ricardo, Guilherme, Adriana, Luciano, Rafael; e à geração Y: Marianella, Letícia, Luís Felipe, Danilo, Amanda, Maria Izabel, Fernando, Rodrigo Júnior, Cristiane e Jhonatthan. A convivência diária com vocês tornou essa etapa muito mais prazerosa. Já sinto saudades do Lab.

À Graciela, Guilherme, Adriana e João, pelas amizades especiais. Graci, tornou-se uma irmã muito querida, é meu exemplo de amor ao melhoramento de plantas. Gui, com sua inteligência surpreendente e suas palavras carinhosas, sempre me confortou e motivou quando preciso. Dri, foi e sempre será meu exemplo de foco, garra e dedicação. E o velho Jones, que me acompanha nessa trajetória desde o início da ESALQ, além de ser muito centrado e objetivo, tornou-se meu irmão de coração.

Algumas parcerias de trabalho de curto prazo serão especiais para o resto da vida: Melina Mancini, Augusto Lima e Tábata Bergonci, muito obrigada pela amizade. Assim, como minha estadia em Piracicaba que foi abençoada com a convivência da Bruna Mendes, que desde a graduação acompanha esse sonho no carão do abacate, e da Marina

Duarte (Kapiaoa). À Lu Carlin-Garcia, obrigada pelo carinho sem dimensão.

A todo o pessoal do Departamento de Genética da ESALQ: Léia, Fernando, Macedônio, Seo Natálio, Valdir, Berdan, Maídia, Seo Antônio, Marcos e demais funcionários.

Ao Prof. Luiz Alexandre Peternelli, por me apresentar a ciência de uma maneira fascinadora e encantadora.

Aos anjos terrestres que me guiaram nessa trajetória, Joana e Leonardo, por me acolherem como filha; Renata Rueda, pelo carinho, Luiz Hercoton, meu pai piracicabano; Denise Perígo, Fernanda Lima, Ana Maria, Nathee Pelinson, Aline Torres e Nataly Carbonaro, pelo segredo da amizade.

Finalmente, a todos os amigos e familiares que me ajudaram direta ou indiretamente para que esse sonho se tornasse realidade. Muito obrigada!

EPIGRAFE

*“Tudo tem o seu tempo determinado,
e há tempo para todo o propósito debaixo do céu.”
Eclesiastes 3:1*

SUMÁRIO

1	Introdução	11
2	Considerações Finais	17
	Referências	19

RESUMO

Mapeamento associativo para múltiplos ambientes e múltiplos locos, visando tolerância à seca em milho

A seca é um dos estresses abióticos mais importantes na cultura do milho, o qual ocasiona reduções significativas na produção de grãos. A arquitetura genética da tolerância à seca é complexa, fazendo-se necessária a melhor compreensão desse caráter. Estudos envolvendo mapeamento associativo são úteis por explorarem a variação genética de caracteres quantitativos e, adicionalmente, levam em conta informações acerca de genótipos, ambientes e interações genótipo por ambiente ($G \times E$). Ao considerar efeitos de $G \times E$ em modelos de mapeamento associativo há possibilidade de identificar regiões no genoma associadas às condições e ambientes específicos. Este trabalho teve como objetivo detectar associações relacionadas à tolerância à seca em milho por meio de um modelo de mapeamento associativo para múltiplos ambientes e múltiplos locos, o qual permitiu distinguir associações com efeitos ambiente-específico daquelas com efeitos principais e de interação associação por ambiente (QEI). O painel associativo foi composto por 190 linhagens, classificadas de acordo com os grupos heteróticos quanto ao tipo de grão. Marcadores SNPs (~500k) foram utilizados para a genotipagem do painel associativo. Duas linhagens (L228-3 e L3) foram usadas como testadores comuns e os híbridos obtidos foram avaliados em duas localidades (Janaúba-MG e Teresina-PI), dois anos agrícolas (2010 e 2011), sob duas condições de tratamento (irrigado e não irrigado). Ao total, consideraram-se seis caracteres: peso de grãos, intervalo de florescimento, florescimento feminino e masculino, altura de planta e de espiga. Consideraram-se dois grupos de mapeamento, agrupados de acordo com os testadores utilizados. SNPs foram úteis para testar associações ao longo do genoma do milho e investigar o relacionamento genético entre indivíduos. O modelo de mapeamento associativo, com inclusão de informações sobre interação $G \times E$, detectou o total de 179 associações, e o maior número de associações foram relacionadas aos caracteres de florescimento. A maioria das associações (168) apresentaram QEI significativo, sendo que o tamanho e a magnitude desses efeitos distinguiram-se de acordo com o ambiente em avaliação. Apenas o caráter florescimento feminino não apresentou associações com efeitos estáveis ao longo dos ambientes em estudo. A detecção de algumas associações em posições próximas do genoma evidenciam possíveis efeitos de pleiotropia. Algumas associações foram co-localizadas em regiões do genoma do milho relacionadas à tolerância à seca, sendo que algumas dessas associações estavam envolvidas a fatores pertencentes às vias metabólicas de interesse. O presente estudo forneceu informações úteis para a compreensão da base genética da tolerância à seca em milho sob os ambientes específicos em avaliação.

Palavras-chave: *Zea mays*, GWAS, $G \times E$, Tolerância à seca

ABSTRACT

Multi-environment-multi-locus association mapping for drought tolerance in maize

Drought is a severe stress factor in maize production and causes significant reduction in grain yield. Genetic architecture of drought tolerance is complex and a better understanding of this trait is required. Association mapping studies are useful to explore quantitative traits and simultaneously account for genetic backgrounds including genotype, environment and genotype-by-environment ($G \times E$) interactions. By accounting for $G \times E$ into association mapping models it is possible to identify regions associated with specific environment and conditions. The main goal of this study was detect significant associations related to drought tolerance in maize via multi-environment-multi-locus association mapping model, distinguishing information about specific-environment effects from main and association-by-environment (QEI) effects. Our association panel was composed by 190 inbred lines classified according to heterotic groups. The panel was genotyped with ~500K SNPs. Two inbred lines (L228-3 and L3) were used as common testers and testcrosses were assessed in two locations (Janaúba-MG e Teresina-PI), two years (2010 and 2011), under two treatment conditions (well-watered and water-stressed). A total of six traits were evaluated including grain yield, anthesis-silking interval, female and male flowering time, plant and ear height. Two mapping groups were considered, grouped by common testers. SNPs were used to test significant association along the maize genome and also to account for population structure and relatedness coefficient. Our mapping model detected a total of 179 associations and the highest number of associations were related to the flowering time measures. The most associations (168) showed significant QEI and the size and magnitude of those effects were distinguished by environment conditions. Only female flowering trait did not show stable effects across all environments. Mapped associations in nearby positions indicate plausible pleiotropy effects. Some associations were co-located in maize genome regions related to metabolic pathway factors. Our study support the detection of significant associations along the maize genome and contributes for understanding of the genetic basis of drought tolerance in maize.

Keywords: *Zea mays*, GWAS, $G \times E$, Drought tolerance

n

1 INTRODUÇÃO

O milho é uma das culturas mais importantes produzidas mundialmente. Além de ser fonte principal para produção de amido, é matéria-prima de produtos como óleo, açúcar, etanol, entre outros, constituindo-se como base para a alimentação humana e animal (XUE ET AL., 2013). Com o aumento da população e, consequentemente, do consumo de alimentos e produtos derivados do milho, há demanda por maior produção por unidade de área plantada. O aumento da sua produtividade em regiões tropicais, como o Brasil, é um desafio. Isto porque, além da produção ser influenciada por vários tipos de estresses bióticos e abióticos, sabe-se que grande parte dos caracteres de importância agrônômica e econômica do milho possui natureza quantitativa e sofre grande influência ambiental. Diante deste cenário, o melhoramento de plantas, mostra-se como uma ferramenta útil pois, além de possibilitar maior informação sobre a natureza complexa dos caracteres em estudo, permite a seleção de genótipos superiores sob condições específicas em avaliação, contribuindo para a obtenção de ganhos em produtividade dessa cultura.

Botanicamente, o milho pertence à família Poaceae, subfamília Panicoideae, tribo Maydeae, gênero *Zea*, espécie *Zea Mays* L. É uma espécie alógama e diploide, com o número de cromossomos $x=10$, $2n=2x=20$. O tamanho de seu genoma é de aproximadamente 2500 megabases (Mb) e sua característica notável é a grande diversidade molecular, caracterizada principalmente pela existência de elementos transponíveis (LISCH, 2012). Muito embora o milho seja uma das espécies mais estudadas, sua origem e domesticação não são muito bem definidas (BOREM ET AL., 2015). No entanto, a abundância de diversidade existente permite que novos conhecimentos sejam incorporados constantemente na compreensão do processo de domesticação bem como no aperfeiçoamento dos programas de melhoramento dessa espécie (BUCKLER ET AL., 2006).

De modo geral, as etapas dos programas de melhoramento de milho baseiam-se no desenvolvimento de linhagens a partir de populações parentais, avaliação de tais linhagens em cruzamento, seguido da identificação e seleção de híbridos superiores. Para tanto, realiza-se seleção fenotípica dos caracteres de importância comercial e agrônômica, os quais são controlados por muitos genes e são influenciados pelo ambiente (HALLAUER and CARENA, 2009).

A maioria dos programas de melhoramento de milho para regiões de clima subtropical e tropical enfrentam o desafio de produzir materiais tolerantes à seca, uma vez que nessas regiões tais materiais desenvolvem-se sob condições não favoráveis, como altas temperaturas e déficit hídrico (FARFAN ET AL., 2015). De fato, a seca é um dos maiores estresses abióticos na produção do milho, o qual ocasiona perdas significativas na produtividade de grãos. Especificamente no Brasil, nos últimos anos, a precipitação anual ficou abaixo da média na maior parte das regiões produtoras no país (INMET, 2015). Consequente ao aquecimento global e aos baixos níveis de precipitação, problemas no abastecimento de água em grandes centros urbanos bem como a utilização desse bem pelo setor elétrico fazem com que a disponibilidade de água ao setor agrícola seja menor. Além disso, uma grande parcela dos produtores não possuem a irrigação como tecnologia disponível para suprir as condições de déficit hídrico na agricultura. Diante desses fatos, nota-se a demanda por materiais que possuam alta eficiência no uso da água e, nesse contexto, o desenvolvimento de materiais tolerantes à seca tem se tornado prioridade dos programas de melhoramento de milho (RIBAUT ET AL., 2009).

A tolerância à seca é um caráter bastante complexo, difícil de ser quantificado e melhorado, e, além de ser

regulado por milhares de genes (CAMPOS *ET AL.*, 2004), está associado à várias vias metabólicas (RENGEL *ET AL.*, 2012). Adicionalmente, há níveis diferentes de tolerância, os quais dependem da quantidade, intensidade e tempo decorrido do estresse (BRUCE *ET AL.*, 2002). Em milho, o estresse hídrico é um fator limitante principalmente durante o período de florescimento e enchimento de grãos, sendo que maiores perdas na produtividade são observadas quando o estresse ocorre nesses estágios de desenvolvimento (BOLANOS and EDMAN, 1996). Alterações significativas como redução na altura da planta, peso de espiga e área foliar; maior incidência de tombamento de plantas; aborto de embriões; senescência prematura das folhas; redução da atividade fotossintética; entre outras, são também observadas (RIBAUT *ET AL.*, 2009). Devido à complexidade genética, para a obtenção de ganhos no melhoramento objetivando tolerância à seca, faz-se necessários estudos que permitam o entendimento da arquitetura genética desse caráter e ainda levem em consideração informações acerca de genótipos, ambientes e interações genótipo por ambiente ($G \times E$) (XUE *ET AL.*, 2013). Isto porque espera-se que a performance diferencial de genótipos tolerantes em relação aos não-tolerantes ao estresse produza interação $G \times E$; portanto, uma análise adequada dessa interação é pré-requisito para o sucesso dos programas de melhoramento (VAN EEUWIJK *ET AL.*, 2007).

A compreensão dos mecanismos genéticos da tolerância à seca em milho tem sido estudada corriqueiramente por meio de marcadores moleculares. Além de acessarem a variação do genoma do milho via diversidade genética (BE-TRAN *ET AL.*, 2003; LABORDA *ET AL.*, 2005; SCHNABLE *ET AL.*, 2009; MARON *ET AL.*, 2013) e construção de mapas genéticos (SIBOV *ET AL.*, 2003; COURTIAL *ET AL.*, 2013; AZEVEDO *ET AL.*, 2015; KHANAL *ET AL.*, 2015), por exemplo, estudos envolvendo o uso de marcadores moleculares exploram associações entre genótipos e caracteres fenotípicos, permitindo a identificação de genes funcionais ou marcadores ligados a genes que se expressam durante o estresse hídrico. Vários estudos com o objetivo de detectar associações relacionadas à tolerância à seca em milho foram conduzidos por meio do mapeamento de locos para caracteres quantitativos (do inglês, *quantitative trait loci* - QTL), considerando populações biparentais e técnicas diferentes de mapeamento (RIBAUT *ET AL.*, 1996, 1997; TUBEROSA *ET AL.*, 2002; LU *ET AL.*, 2010; ALMEIDA *ET AL.*, 2013). No entanto, as associações identificadas nessas condições geralmente não podem ser extrapoladas para outros tipos de populações e não são capazes de acessar a diversidade alélica existente ao longo de todo o genoma (XUE *ET AL.*, 2013). Por outro lado, estudos envolvendo mapeamento associativo, além de serem uma ferramenta útil para a detecção de associações relacionadas às características de interesse, apresentam vantagens quando comparados com o mapeamento de QTLs, como por exemplo *i*) aumento da resolução e precisão de mapeamento para a localização de associações, *ii*) maior variação alélica, *iii*) redução no tempo de pesquisa e consequente redução nos custos, visto que não há necessidade do desenvolvimento de populações experimentais (ZHU *ET AL.*, 2008). Outra opção que tem se tornado comum é realizar simultaneamente mapeamento associativo e mapeamento de QTLs, combinando as vantagens de ambos (YU *ET AL.*, 2008; McMULLEN *ET AL.*, 2009).

Em linhas gerais, o mapeamento associativo, também conhecido como mapeamento por desequilíbrio de ligação, consiste na identificação de locos que controlam caracteres quantitativos, de modo que a associação detectada possa ser atribuída ao desequilíbrio de ligação entre marcadores e polimorfismos funcionais em um conjunto de genótipos (YU *ET AL.*, 2006). Em estudos de associação o foco está na detecção de mutações antigas que se mantiveram ao longo da população. Uma vez que neste último cenário há aumento do número de gerações de recombinação, regiões

muito menores do genoma contendo a mutação original serão herdadas e então, a associação de marcadores com o caráter de interesse é detectada mais precisamente (ZHU *ET AL.*, 2008). Por explorar eventos de recombinação evolutivos e históricos em nível de população, o mapeamento associativo tem se tornado uma ferramenta alternativa para o entendimento da variação de características quantitativas (NORDBORG and TAVARÉ, 2002). Estudos de mapeamento associativo são realizados considerando dois interesses principais, sendo que o primeiro deles envolve o estudo de um gene candidato, enquanto que o segundo considera o estudo do genoma como um todo (do inglês, *genome-wide association study* - GWAS). Esta segunda abordagem foi facilitada pela atual disponibilidade de plataformas de sequenciamento de nova geração. Muito embora o mapeamento associativo de um gene candidato seja útil para estabelecer associações entre genes específicos e determinados caracteres complexos, a maioria dos estudos de associação em milho relacionados à seca foram realizados considerando a abordagem de GWAS (HAO *ET AL.*, 2011; LIU *ET AL.*, 2013; XUE *ET AL.*, 2013; THIRUNAVUKKARASU *ET AL.*, 2014; FARFAN *ET AL.*, 2015).

Em estudos de associação, um dos grandes fatores de confundimento é o grau de relacionamento genético entre indivíduos, o qual pode ocasionar detecção de associações falso-positivas. Por esse motivo, a fim de controlar tais efeitos, uma prática comum é utilizar a estrutura populacional e o parentesco como fatores dentro de um modelo de associação (PRITCHARD *ET AL.*, 2000). Neste sentido, para a maioria das estratégias de mapeamento associativo, primeiramente verifica-se a informação de parentesco e a existência de estrutura populacional entre indivíduos, usando-as eventualmente para inferência de correlações genéticas em um modelo de mapeamento. Assim, ao considerar tais fatores, a associação marcador-caráter (do inglês, *marker trait association* - MTA) é testada dentro de cada subgrupo, removendo as correlações genéticas devido ao efeito de grupos (MALOSETTI *ET AL.*, 2007). Investigar o grau de relacionamento genético em milho é fundamental para que as associações detectadas por meio do mapeamento associativo sejam, de fato, devidas ao desequilíbrio de ligação. Isto porque sabe-se que a maioria dos programas de melhoramento estabelecidos ao redor do mundo trabalham com o desenvolvimento de materiais a partir da exploração de grupos heteróticos distintos (UNTERSEER *ET AL.*, 2014) e, uma característica marcante de grande parte dos germoplasmas desses programas, é a subdivisão dos materiais vegetais em grupos heteróticos. Além disso, espera-se que parte dos materiais que compõem o germoplasma de milho apresentem algum grau de relacionamento genético, seja pelo fato de coincidências de suas origens geográficas, locais de adaptações ou ainda histórias de melhoramento (COOPER *ET AL.*, 2014).

Uma ferramenta útil e efetiva que permite a incorporação nas análises da informação de estrutura de população e parentesco em estudos de mapeamento associativo são os modelos mistos lineares, mais popularmente conhecidos por modelos mistos (ZHOU and STEPHENS, 2012). Ao contrário dos modelos lineares tradicionais, os quais assumem independência estatística entre observações, ao considerar o GWAS via modelagem mista é possível incorporar a correlação ou dependência existente entre indivíduos por meio de informações de estrutura populacional e parentesco (HOFFMAN, 2013). Geralmente, a correlação entre indivíduos é modelada por meio da regressão dos fenótipos em função de uma matriz de marcadores moleculares, usando a estrutura populacional, representada por exemplo, por *scores* de componentes principais, como covariável fixa do modelo genético estatístico. Já o coeficiente de parentesco, geralmente estimado com base em marcadores moleculares (do inglês, *realized relationship matrix*) é incluído na parte

aleatória do modelo, na forma de matriz de variâncias e covariâncias (VCOV), a qual representa a dependência entre os indivíduos (KANG *ET AL.*, 2010). Adicionalmente, os modelos mistos são úteis para a avaliação de dados fenotípicos de programas de melhoramento, os quais muitas vezes são também utilizados em estudos de associação. Isto porque dados fenotípicos de programas de melhoramento geralmente são provenientes de ensaios realizados em vários ambientes (do inglês, *multi-environment trials* - MET) (VAN EEUWIJK and KROONENBERG, 1998), avaliados ao longo de localidades e anos agrícolas diferentes e, apresentam corriqueiramente grau elevado de desbalanceamento (MÖHRING and PIEPHO, 2009). Neste contexto, ao usar a abordagem de modelos mistos, é possível tratar o desbalanceamento de dados e, ainda, inferir correlações entre localidades e anos agrícolas bem como a ocorrência de interações $G \times E$ por meio da estruturação de matrizes de VCOV apropriadas (PIEPHO, 1997; VAN EEUWIJK and KROONENBERG, 1998; CULLIS *ET AL.*, 1998; SMITH *ET AL.*, 2001, 2005).

Existem inúmeros procedimentos para a realização de uma análise envolvendo modelos mistos, sendo que tais procedimentos variam de acordo com o objetivo do estudo. Segundo PIEPHO *ET AL.* (2012), especificamente em estudos envolvendo mapeamento associativo, é conveniente considerar a modelagem mista de acordo com a abordagem em dois estágios (do inglês, *two-stage approach*) (SMITH *ET AL.*, 2001) pelo fato de ser mais simplificada e computacionalmente mais rápida. De fato, a maioria dos trabalhos de GWAS em milho para caracteres relacionados à seca foram realizados considerando a abordagem de modelagem mista em dois estágios (HAO *ET AL.*, 2011; LIU *ET AL.*, 2013; XUE *ET AL.*, 2013; THIRUNAVUKKARASU *ET AL.*, 2014; FARFAN *ET AL.*, 2015). Resumidamente, em tais trabalhos, no primeiro estágio realizou-se a análise dos caracteres fenotípicos, avaliados ao longo de localidades, anos agrícolas e condições de tratamento (estresse ou não estresse hídrico) diferentes, por meio de um modelo misto apropriado. Em algumas situações, houve modelagem de variâncias e covariâncias genéticas e ou residuais e então, médias ajustadas foram obtidas para cada ambiente, ou conjuntamente de acordo com as condições de tratamento, ou considerando uma análise conjunta de todos os ambientes simultaneamente (característica de análise envolvendo dados do tipo MET). Então, no segundo estágio da análise, médias ajustadas foram utilizadas como informação dos fenótipos dos indivíduos dentro do modelo de mapeamento associativo, juntamente com a inclusão de informações sobre o grau de relacionamento entre eles, seja pela incorporação de estrutura de população e/ou do parentesco.

Muito embora a maioria dos trabalhos de GWAS em milho visando tolerância à seca envolvam caracteres fenotípicos de natureza complexa, avaliados em condições que remetem a dados de MET, tradicionalmente, ajusta-se um modelo misto apropriado para cada caráter e ambiente e, então, realiza-se o GWAS para cada ambiente, desconSIDERANDO informações que dizem respeito à interação $G \times E$ e QTL (ou associação) por ambiente (do inglês, *QTL by environment interaction* - QEI). Apenas FARFAN *ET AL.* (2015) relatam explicitamente a exploração da interação $G \times E$ na avaliação dos caracteres fenotípicos em milho relacionados à seca, via modelagem mista de uma matriz de VCOV de um modelo não-estruturado, sendo que a exploração da QEI foi realizada de maneira informal, por meio da análise de gráficos e tabelas. MALOSETTI *ET AL.* (2004), afirmam que embora a interação $G \times E$ possua papel fundamental em modelos de respostas fenotípicas, poucos trabalhos reúnem esforços para a incorporação da base genética da interação $G \times E$, expressa por meio da QEI. BOER *ET AL.* (2007) sugerem que testes explícitos devem ser realizados para a presença, magnitude e forma da interação $G \times E$, com o intuito de compreender a arquitetura genética dos

caracteres complexos. Adicionalmente, VAN EEUWIJK *ET AL.* (2007) endossam que uma opção atrativa para estudos da base genética da tolerância ao estresse abiótico e interação $G \times E$ é o desenvolvimento de uma abordagem que trate simultaneamente da modelagem da interação $G \times E$ e do mapeamento de QTLs.

No nosso conhecimento, o único trabalho de GWAS que considera a incorporação da interação $G \times E$ no modelo de mapeamento associativo foi desenvolvido por GUTIÉRREZ *ET AL.* (2015). Neste trabalho, considerou-se um modelo misto no contexto de GWAS para múltiplos ambientes e múltiplos QTLs, para caracteres relacionados à resistência à doença em cevada. Em milho, nenhum trabalho de mapeamento associativo relata o estudo da interação $G \times E$ dentro do modelo de associação. No entanto, considerar tais aspectos no GWAS, visando tolerância à seca em milho é importante, não apenas pelo fato dos caracteres relacionados ao estresse hídrico apresentarem interação $G \times E$, como comentado anteriormente, mas também por ser possível distinguir associações com efeitos principais daquelas com efeitos ambiente-específico. Tais aspectos, além de auxiliarem na compreensão da base genética dos caracteres em estudos, são aplicáveis aos programas de melhoramento, uma vez que torna-se possível a identificação de regiões associadas à genes expressos em condições e ambientes específicos, bem como o uso de tais informações, na seleção de genótipos superiores por meio da seleção assistida por marcadores, dentre outras aplicações.

Neste sentido, o objetivo do presente trabalho foi o desenvolvimento de um modelo misto para mapeamento associativo considerando múltiplos ambientes e múltiplos locos, visando estudar a arquitetura genética de caracteres relacionados à tolerância à seca para germoplasma brasileiro de milho. Para tanto, considerou-se o estudo da interação $G \times E$ por meio da modelagem de uma matriz de VCOV para os ambientes em avaliação, bem como a realização de testes explícitos para detecção de associações com efeitos principais ou ambiente-específicos.

2 CONSIDERAÇÕES FINAIS

Estudos de mapeamento associativo são conduzidos corriqueiramente considerando cada vez mais maiores quantidades de dados genotípicos e fenotípicos. A motivação principal para tal fato é a possibilidade de explicar a variação fenotípica em termos de mudanças na sequência de DNA (polimorfismos). Este é o primeiro estudo de mapeamento associativo visando tolerância à seca em milho realizado para múltiplos ambientes e múltiplos locos simultaneamente, o qual incorporou informações de interação $G \times E$ e QEI diretamente no modelo de mapeamento.

Ao fazer uso de ferramentas baseadas na abordagem de modelos mistos, aumentou-se o poder das análises de mapeamento por considerar modelos de estruturas de VCOV específicos, os quais apresentaram variâncias heterogêneas e existência de correlações entre as observações. Adicionalmente, a inclusão de informações sobre relacionamento genético entre os indivíduos em estudo, controlou possíveis fatores de confundimento, intrínsecos à estudos de associação.

A exploração da estrutura populacional e do parentesco possibilitou inferir que a maioria dos indivíduos pertencentes ao painel associativo da Embrapa Milho e Sorgo apresentam-se como um grupo intermediário aos grupos heteróticos *Dent* e *Flint*, o que é justificável pela recente história de melhoramento desses materiais. Assim, a subdivisão em dois grupos de mapeamento associativo, organizados de acordo com os testadores, é de grande valia para a compreensão da arquitetura genética da tolerância à seca dentro de cada grupo de mapeamento.

Adicionalmente, é válido mencionar que a inclusão de informações sobre interação $G \times E$ e QEI, demonstrou a importância de considerar tais fatores no presente mapeamento associativo. Para todos os caracteres e grupos de mapeamento em estudo, um número grande de associações foram encontradas e, notou-se a expressão significativa da interação genótipo por ambiente ou associação por ambiente. De fato, os experimentos de campos, os quais incluíram diferentes localidades, anos agrícolas e condições de tratamento, foram delineados para que tais interações fossem expressas. No entanto, essas condições representam situações enfrentadas corriqueiramente pelos melhoristas ao avaliarem certos materiais em diferentes ambientes.

Diante desse cenário é possível concluir que os resultados aqui obtidos são importantes e servem como base não apenas para estudos futuros que objetivem a compreensão de vias metabólicas e mecanismos genéticos do estresse hídrico, mas também para auxiliar os melhoristas de milho, seja por meio da seleção assistida por marcadores ou em etapas do melhoramento convencional.

REFERÊNCIAS

- ALMEIDA, G. D., D. MAKUMBI, C. MAGOROKOSHO, S. NAIR, A. BORÉM, J.-M. RIBAUT, M. BÄNZIGER, B. M. PRASANNA, J. CROSSA, and R. BABU, 2013 Qtl mapping in three tropical maize populations reveals a set of constitutive and adaptive genomic regions for drought tolerance. *Theoretical and Applied Genetics* **126**: 583–600.
- AZEVEDO, G. C., A. CHEAVEGATTI-GIANOTTO, B. F. NEGRI, B. HUFNAGEL, L. D. C. E SILVA, J. V. MAGALHAES, A. A. F. GARCIA, U. G. LANA, S. M. DE SOUSA, and C. T. GUIMARAES, 2015 Multiple interval qtl mapping and searching for pstol1 homologs associated with root morphology, biomass accumulation and phosphorus content in maize seedlings under low-p. *BMC plant biology* **15**: 1.
- BETRAN, F., J. RIBAUT, D. BECK, and D. G. DE LEON, 2003 Genetic diversity, specific combining ability, and heterosis in tropical maize under stress and nonstress environments. *Crop Science* **43**: 797–806.
- BOER, M. P., D. WRIGHT, L. FENG, D. W. PODLICH, L. LUO, M. COOPER, and F. A. VAN EEUWIJK, 2007 A mixed-model quantitative trait loci (qtl) analysis for multiple-environment trial data using environmental covariables for qtl-by-environment interactions, with an example in maize. *Genetics* **177**: 1801–1813.
- BOLANOS, J. and G. EDMEADES, 1996 The importance of the anthesis-silking interval in breeding for drought tolerance in tropical maize. *Field Crops Research* **48**: 65–80.
- BOREM, A., J. C. C. GALVAO, and M. A. PIMENTEL, 2015 Milho: do plantio a colheita. UFV.
- BRUCE, W. B., G. O. EDMEADES, and T. C. BARKER, 2002 Molecular and physiological approaches to maize improvement for drought tolerance. *Journal of experimental botany* **53**: 13–25.
- BUCKLER, E. S., B. S. GAUT, and M. D. McMULLEN, 2006 Molecular and functional diversity of maize. *Current opinion in plant biology* **9**: 172–176.
- CAMPOS, H., M. COOPER, J. HABBEN, G. EDMEADES, and J. SCHUSSLER, 2004 Improving drought tolerance in maize: a view from industry. *Field Crops Research* **90**: 19–34.
- COOPER, M., C. GHO, R. LEAFGREN, T. TANG, and C. MESSINA, 2014 Breeding drought-tolerant maize hybrids for the us corn-belt: discovery to product. *Journal of Experimental Botany* p. eru064.
- COURTIAL, A., J. THOMAS, M. REYMOND, V. MÉCHIN, J. GRIMA-PETTENATI, and Y. BARRIÈRE, 2013 Targeted linkage map densification to improve cell wall related qtl detection and interpretation in maize. *Theoretical and Applied Genetics* **126**: 1151–1165.
- CULLIS, B., B. GOGEL, A. VERBYLA, and R. THOMPSON, 1998 Spatial analysis of multi-environment early generation variety trials. *Biometrics* pp. 1–18.
- FARFAN, I. D. B., G. N. DE LA FUENTE, S. C. MURRAY, T. ISAKEIT, P.-C. HUANG, M. WARBURTON, P. WILLIAMS, G. L. WINDHAM, and M. KOLOMIETS, 2015 Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. *PloS one* **10**: e0117737.

- GUTIÉRREZ, L., S. GERMÁN, S. PEREYRA, P. M. HAYES, C. A. PÉREZ, F. CAPETTINI, A. LOCATELLI, N. M. BERBERIAN, E. E. FALCONI, R. ESTRADA, ET AL., 2015 Multi-environment multi-qtI association mapping identifies disease resistance qtI in barley germplasm from latin america. *Theoretical and Applied Genetics* **128**: 501–516.
- HALLAUER, A. R. and M. J. CARENA, 2009 Maize. Springer.
- HAO, Z., X. LI, C. XIE, J. WENG, M. LI, D. ZHANG, X. LIANG, L. LIU, S. LIU, and S. ZHANG, 2011 Identification of functional genetic variations underlying drought tolerance in maize using snp markers. *Journal of integrative plant biology* **53**: 641–652.
- HOFFMAN, G. E., 2013 Correcting for population structure and kinship using the linear mixed model: theory and extensions. *PloS one* **8**: e75707.
- INMET, 2015 Plano Estrategico 2014-2015.
- KANG, H. M., J. H. SUL, S. K. SERVICE, N. A. ZAITLEN, S.-Y. KONG, N. B. FREIMER, C. SABATTI, E. ESKIN, ET AL., 2010 Variance component model to account for sample structure in genome-wide association studies. *Nature genetics* **42**: 348–354.
- KHANAL, R., A. NAVABI, and L. LUKENS, 2015 Linkage map construction and quantitative trait loci (qtI) mapping using intermated vs. selfed recombinant inbred maize line (zea mays l.). *Canadian Journal of Plant Science* **95**: 1133–1144.
- LABORDA, P., K. OLIVEIRA, A. GARCIA, M. PATERNIANI, and A. DE SOUZA, 2005 Tropical maize germplasm: what can we say about its genetic diversity in the light of molecular markers? *Theoretical and Applied Genetics* **111**: 1288–1299.
- LISCH, D., 2012 Regulation of transposable elements in maize. *Current opinion in plant biology* **15**: 511–516.
- LIU, S., X. WANG, H. WANG, H. XIN, X. YANG, J. YAN, J. LI, L.-S. P. TRAN, K. SHINOZAKI, K. YAMAGUCHI-SHINOZAKI, ET AL., 2013 Genome-wide analysis of zmdreb genes and their association with natural variation in drought tolerance at seedling stage of zea mays l .
- LU, Y., S. ZHANG, T. SHAH, C. XIE, Z. HAO, X. LI, M. FARKHARI, J.-M. RIBAUT, M. CAO, T. RONG, ET AL., 2010 Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proceedings of the National Academy of Sciences* **107**: 19585–19590.
- MALOSETTI, M., C. G. VAN DER LINDEN, B. VOSMAN, and F. A. VAN EEUWIJK, 2007 A mixed-model approach to association mapping using pedigree information with an illustration of resistance to phytophthora infestans in potato. *Genetics* **175**: 879–889.
- MALOSETTI, M., J. VOLTAS, I. ROMAGOSA, S. ULLRICH, and F. VAN EEUWIJK, 2004 Mixed models including environmental covariables for studying qtI by environment interaction. *Euphytica* **137**: 139–145.
- MARON, L. G., C. T. GUIMARÃES, M. KIRST, P. S. ALBERT, J. A. BIRCHLER, P. J. BRADBURY, E. S. BUCKLER, A. E. COLUCCIO, T. V. DANILOVA, D. KUDRNA, ET AL., 2013 Aluminum tolerance in maize is associated with higher mate1 gene copy number. *Proceedings of the National Academy of Sciences* **110**: 5241–5246.

- McMULLEN, M. D., S. KRESOVICH, H. S. VILLEDA, P. BRADBURY, H. LI, Q. SUN, S. FLINT-GARCIA, J. THORNSBERRY, C. ACHARYA, C. BOTTOMS, ET AL., 2009 Genetic properties of the maize nested association mapping population. *Science* **325**: 737–740.
- MÖHRING, J. and H.-P. PIEPHO, 2009 Comparison of weighting in two-stage analysis of plant breeding trials. *Crop science* **49**: 1977–1988.
- NORDBORG, M. and S. TAVARÉ, 2002 Linkage disequilibrium: what history has to tell us. *TRENDS in Genetics* **18**: 83–90.
- PIEPHO, H. P., 1997 Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* pp. 761–766.
- PIEPHO, H. P., J. MÖHRING, T. SCHULZ-STREECK, and J. O. OGUTU, 2012 A stage-wise approach for the analysis of multi-environment trials. *Biometrical Journal* **54**: 844–860.
- PRITCHARD, J. K., M. STEPHENS, N. A. ROSENBERG, and P. DONNELLY, 2000 Association mapping in structured populations. *The American Journal of Human Genetics* **67**: 170–181.
- RENGEL, D., S. ARRIBAT, P. MAURY, M.-L. MARTIN-MAGNIETTE, T. HOURLIER, M. LAPORTE, D. VARÈS, S. CARRÈRE, P. GRIEU, S. BALZERGUE, ET AL., 2012 A gene-phenotype network based on genetic variability for drought responses reveals key physiological processes in controlled and natural environments. *PloS one* **7**: e45249.
- RIBAUT, J.-M., J. BETRAN, P. MONNEVEUX, and T. SETTER, 2009 Drought tolerance in maize. In Handbook of maize: its biology, pp. 311–344, Springer.
- RIBAUT, J.-M., D. HOISINGTON, J. DEUTSCH, C. JIANG, and D. GONZALEZ-DE LEON, 1996 Identification of quantitative trait loci under drought conditions in tropical maize. 1. flowering parameters and the anthesis-silking interval. *Theoretical and Applied Genetics* **92**: 905–914.
- RIBAUT, J.-M., C. JIANG, D. GONZALEZ-DE LEON, G. EDMÉADES, and D. HOISINGTON, 1997 Identification of quantitative trait loci under drought conditions in tropical maize. 2. yield components and marker-assisted selection strategies. *Theoretical and Applied Genetics* **94**: 887–896.
- SCHNABLE, P. S., D. WARE, R. S. FULTON, J. C. STEIN, F. WEI, S. PASTERNAK, C. LIANG, J. ZHANG, L. FULTON, T. A. GRAVES, ET AL., 2009 The b73 maize genome: complexity, diversity, and dynamics. *science* **326**: 1112–1115.
- SIBOV, S. T., C. LOPES DE SOUZA JR, A. A. GARCIA, A. GARCIA, A. R. SILVA, C. A. MANGOLIN, L. L. BENCHIMOL, and A. P. DE SOUZA, 2003 Molecular mapping in tropical maize (*zea mays* l.) using microsatellite markers. 1. map construction and localization of loci showing distorted segregation. *Hereditas* **139**: 96–106.
- SMITH, A., B. CULLIS, and R. THOMPSON, 2001 Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* pp. 1138–1147.

- SMITH, A., B. R. CULLIS, and R. THOMPSON, 2005 The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science* **143**: 449–462.
- THIRUNAVUKKARASU, N., F. HOSSAIN, K. ARORA, R. SHARMA, K. SHIRIGA, S. MITTAL, S. MOHAN, P. M. NAMRATHA, S. DOGGA, T. S. RANI, *ET AL.*, 2014 Functional mechanisms of drought tolerance in subtropical maize (*zea mays* l.) identified using genome-wide association mapping. *BMC genomics* **15**: 1182.
- TUBEROSA, R., S. SALVI, M. C. SANGUINETI, P. LANDI, M. MACCAFERRI, and S. CONTI, 2002 Mapping qtls regulating morpho-physiological traits and yield: Case studies, shortcomings and perspectives in drought-stressed maize. *Annals of Botany* **89**: 941–963.
- UNTERSEER, S., E. BAUER, G. HABERER, M. SEIDEL, C. KNAAK, M. OUZUNOVA, T. MEITINGER, T. M. STROM, R. FRIES, H. PAUSCH, *ET AL.*, 2014 A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k snp genotyping array. *BMC genomics* **15**: 1.
- VAN EEUWIJK, F. A. and P. M. KROONENBERG, 1998 Multiplicative models for interaction in three-way anova, with applications to plant breeding. *Biometrics* pp. 1315–1333.
- VAN EEUWIJK, F. A., M. MALOSETTI, and M. P. BOER, 2007 Modelling the genetic basis of response curves underlying genotype x environment interaction. *Frontis* **21**: 113–124.
- XUE, Y., M. L. WARBURTON, M. SAWKINS, X. ZHANG, T. SETTER, Y. XU, P. GRUDLOYMA, J. GETHI, J.-M. RIBAUT, W. LI, *ET AL.*, 2013 Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. *Theoretical and applied genetics* **126**: 2587–2596.
- YU, J., J. B. HOLLAND, M. D. McMULLEN, and E. S. BUCKLER, 2008 Genetic design and statistical power of nested association mapping in maize. *Genetics* **178**: 539–551.
- YU, J., G. PRESSOIR, W. H. BRIGGS, I. V. BI, M. YAMASAKI, J. F. DOEBLEY, M. D. McMULLEN, B. S. GAUT, D. M. NIELSEN, J. B. HOLLAND, *ET AL.*, 2006 A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics* **38**: 203–208.
- ZHOU, X. and M. STEPHENS, 2012 Genome-wide efficient mixed-model analysis for association studies. *Nature genetics* **44**: 821–824.
- ZHU, C., M. GORE, E. S. BUCKLER, and J. YU, 2008 Status and prospects of association mapping in plants. *The plant genome* **1**: 5–20.