Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Associação genômica ampla para caracteres relacionados à
eficiência no uso de nitrogênio em linhagens de milho tropical

Júlia Silva Morosini

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2017

Júlia Silva Morosini
Engenheira Agrônoma

**Associação genômica ampla para caracteres relacionados à eficiência no uso
de nitrogênio em linhagens de milho tropical**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **ROBERTO FRITSCH NETO**

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Genética e
Melhoramento de Plantas

Piracicaba
2017

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

Morosini, Júlia Silva

Associação genômica ampla para caracteres relacionadas à eficiência no uso de nitrogênio em linhagens de milho tropical / Júlia Silva Morosini. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2017.

49 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. MAS 2. Nitrogênio 3. GWAS 4. SNP I. Título

DEDICATÓRIA

*Aos meus pais, Carlos e Olívia, e à
minha irmã, Natália, pelo amor,
incentivo e suporte incondicionais.*

Dedico

AGRADECIMENTOS

À Deus, pela vida e por todas as bênçãos que derrama em minha família.

À Escola Superior de Agricultura “Luiz de Queiroz” pelas valiosas oportunidades e rica formação. Especialmente ao Departamento de Genética.

Ao Prof. Dr. Roberto Fritsche Neto pela oportunidade, exemplo profissional e, principalmente, pela confiança depositada em mim.

Aos Professores Dr. Carlos Alberto Labate, Dra. Claudia Barros Monteiro Vitorello e Dra. Maria Lúcia Carneiro Vieira pela significativa contribuição em minha caminhada acadêmica.

Aos colegas e funcionários do Laboratório de Melhoramento de Plantas Alógamas pela convivência harmoniosa e pelo auxílio fundamental em meu projeto de mestrado.

Aos meus amados pais, Carlos Alberto Morosini e Olívia Maria B. S. Morosini, por tudo o que fazem por nossa família, sempre com tanto empenho e determinação. À minha querida irmã Natália Silva Morosini pelo apoio, amizade e preciosos conselhos.

Ao Fernando Garcia Espolador pelo companheirismo na vida acadêmica e pessoal, pelo carinho e apoio constantes e por estar sempre ao meu lado mostrando-me que sou capaz.

Ao Laboratório Max Feffer de Plantas, onde passei quatro anos muito proveitosos durante a graduação e fiz amigos para a vida toda.

Ao CNPq pela concessão da bolsa de estudos. À FAPESP (processo nº 2013/24135-2) pelo auxílio financeiro para condução do projeto.

A todos aqui não mencionados que, de alguma forma, contribuíram para a realização deste trabalho.

EPÍGRAFE

"Caminhante, não há caminho; faz-se caminho ao andar".

Antônio Machado

SUMÁRIO

RESUMO	7
ABSTRACT	8
1. INTRODUÇÃO	9
1.1. Objetivos	11
1.1.1. Objetivo geral	11
1.1.2. Objetivos específicos	11
2. REVISÃO BIBLIOGRÁFICA	13
2.1. Cenário nacional do cultivo de milho	13
2.2. Melhoramento com foco em eficiência no uso de nitrogênio	13
2.3. Estudos de Associação Genômica Ampla (GWAS)	15
3. MATERIAL E MÉTODOS	17
3.1. Multiplicação das linhagens	17
3.2. Instalação e condução dos experimentos no campo	17
3.3. Fenotipagem	18
3.4. Cálculo do Índice de Tolerância à Baixa disponibilidade de Nitrogênio	19
3.5. Análises genético-estatísticas dos dados fenotípicos	20
3.6. Genotipagem e controle de qualidade dos dados genômicos	22
3.7. Análises de Associação Genômica Ampla (GWAS)	22
3.8. Desequilíbrio de ligação (LD)	23
3.9. Análise de genes candidatos	24
4. RESULTADOS	25
4.1. Análise de Deviance e valores genotípicos	25
4.2. Estrutura da população	27
4.3. Desequilíbrio de ligação no painel de linhagens	28
4.4. Associações dos SNP com as características fenotípicas	28
4.5. Anotação funcional	31
5. DISCUSSÃO	33
5.1. Análise exploratória dos dados	33
5.2. Associação ampla do genoma	34
5.3. Relações genômicas entre genes candidatos	36
6. CONCLUSÃO	39
REFERÊNCIAS	41
ANEXOS	49

RESUMO

Associação genômica ampla para caracteres relacionadas à eficiência no uso de nitrogênio em linhagens de milho tropical

A expansão dos locais e da época de cultivo do milho faz com que uma fração significativa da produção do cereal ocorra em condições de estresses abióticos. Grande parte desse cenário edafoclimático remete ao cultivo do milho safrinha, semeado entre janeiro e março e atualmente responsável por 65% da produção total. Dentre os diversos tipos de estresses abióticos, a deficiência de nitrogênio é comum e crítica nos solos brasileiros. Evidencia-se, portanto, a necessidade do desenvolvimento de genótipos mais eficientes no uso de nitrogênio, resultando em benefícios econômicos e ambientais. Caracteres morfofisiológicos podem auxiliar no processo seletivo de genótipos superiores para essas condições, como a mensuração do sistema radicular e a análise da taxa fotossintética. Nesse contexto, o objetivo do trabalho foi identificar regiões do genoma do milho tropical associadas ao comprimento de raiz, à fluorescência de clorofila e ao índice de resposta da planta ao estresse por nitrogênio. Para isto, foram avaliadas 64 linhagens endogâmicas de milho tropical em baixa e alta disponibilidade nitrogênio no solo e em dois locais de cultivo na Região de Piracicaba-SP, nas safras 2014/15 e 2015/16. Foram considerados o comprimento de raiz, a eficiência fotossintética do fotossistema II e o Índice de Tolerância à Baixa disponibilidade de nitrogênio (ITBN). As linhagens foram genotipadas com a plataforma Affymetrix® Axiom® Maize Genotyping Array de 616.201 marcadores SNP. A qualidade da informação genômica foi controlada pelos procedimentos *Minor Allele Frequency* e *Call Rate* e pela eliminação de heterozigotos. Os valores genotípicos foram preditos por meio de equações de modelos mistos do tipo REML/BLUP. Os dados de marcadores moleculares e fenotípicos foram analisados por Associação Genômica Ampla (GWAS). Verificou-se maior comprimento radicular em condições de baixa disponibilidade de nitrogênio. No total, sete marcadores significativos foram identificados, sendo quatro referentes ao ITBN, dois referentes ao comprimento de raiz em disponibilidade ideal e um referente ao comprimento de raiz em disponibilidade baixa do nutriente. Entre os principais processos biológicos identificados através da anotação funcional, estão o controle e regulação da transcrição, detectado para todos os caracteres avaliados, e a síntese de Guanosina Monofosfato Sintetase (GMP), enzima diretamente envolvida na disponibilização e reciclagem de nitrogênio. Também foi observada coincidência de região cromossômica entre marcadores significativos identificados e QTL potencialmente relacionados com a eficiência no uso de nitrogênio já reportados na literatura. Conclui-se que a técnica GWAS apresenta eficiência na detecção de marcadores associados aos caracteres de interesse, neste caso evidenciando processos e funções celulares relacionados aos diferentes processos da síntese e reciclagem de nitrogênio.

Palavras-chave: MAS; Nitrogênio; GWAS; SNP

ABSTRACT

Genome-wide association for characters related to nitrogen use efficiency in tropical maize lines

The expansion of locals and season period of maize crop makes a significant portion of the cereal production to occur under abiotic stress conditions. Much of this environmental scenario refers to the second growing season maize, sowed between January and March and responsible for 65% of total production currently. Among the different types of abiotic stresses, nitrogen deficiency is common and critical in Brazilian soils. Therefore, it becomes necessary to develop genotypes more efficient in nitrogen use, resulting in economic and environmental benefits. Morphophysiological characters may assist in the selective process of superior genotypes for these conditions, such as root system measurement and photosynthetic ratio analysis. In this context, this project aimed to identify tropical maize genomic regions associated with root morphological character, physiological parameter of photosystem II, and with the plant response index to nitrogen stress. To this end, 64 tropical maize inbreeding lines contrasting for nitrogen use efficiency were evaluated in low and ideal soil nitrogen availability in two cultivation sites in the region of Piracicaba-SP, in seasons 2014/15 and 2015/16. The characters root total length, chlorophyll fluorescence, and Low Nitrogen Tolerance index (LNTI) were considered. The lines were genotyped SNP using the Affymetrix® Axiom® Maize Genotyping Array with 616,201 SNP. The quality of the genomic information was controlled by Minor Allele Frequency and Call Rate procedures and by the elimination of heterozygous loci. The genotypic values were predicted using REML/BLUP mixed model equations. Genome-Wide Association Studies (GWAS) was performed to analyze molecular and phenotypic data. Greater root length under low availability nitrogen conditions was verified. In total, seven significant markers were identified, four referring to LNTI, two referring to root length under ideal nitrogen availability and one referring to root length under low nitrogen availability. Among the main biological processes identified through functional annotation are the transcription control and regulation, detected to all evaluated characters, and the synthesis of Guanosine Monophosphate Synthetase (GMP), enzyme directly involved in the provision and recycling of nitrogen. It was also observed coincidence of chromosomal region between significant markers identified and QTLs potentially related in nitrogen use efficiency previously reported. As conclusion, GWAS technique shows efficiency in the detection of markers associated to the characters in focus, in this evidencing cellular processes and functions associated to the different process of nitrogen synthesis and recycling.

Keywords: MAS; Nitrogen; GWAS; SNP

1. INTRODUÇÃO

A expansão dos locais e da época de cultivo do milho (*Zea mays* L.), impulsionada pela crescente demanda mundial por alimentos e pelas mudanças climáticas, faz com que uma fração significativa da produção do cereal ocorra em áreas marginais. Nestas áreas, o cultivo é conduzido em condições estressantes regidas por fatores abióticos decorrentes da baixa fertilidade natural dos solos, insuficiente aplicação de fertilizantes nitrogenados e ocorrência de seca (GIÁCOMO et al., 2015; MONNEVEUX et al., 2005).

Entre os diversos tipos de estresses abióticos, a deficiência de nitrogênio (N) é comum nos solos brasileiros, afetando negativamente o desenvolvimento das plantas e reduzindo a produtividade (FRITSCHÉ-NETO; BORÉM, 2012). Ademais, a eficiência na adubação nitrogenada geralmente é menor que 50% devido aos processos de perda por volatilização, desnitrificação e lixiviação (GALLAIS; HIREL, 2004). Deste modo, este fator torna-se preponderante na produção de biomassa, de grãos e nos custos de produção do milho (PENG et al., 2013), assim como na segurança alimentar, uma vez que a cultura do milho é a principal consumidora do adubo nitrogenado (IFA, 2013).

Aliado a isso, o cultivo de milho de segunda safra (milho safrinha), o qual ocorre em regiões e épocas com maior frequência de estresses abióticos, demanda fatores mais complexos para o aumento de produtividade (DUARTE; KAPPES, 2015). Ainda assim, este cultivo apresenta taxa de crescimento de 10% ao ano, enquanto o cultivo de primeira safra cresce cerca de 2% ao ano (GALVÃO et al., 2015). Apesar de apresentar menor custo total em relação ao milho de primeira safra, a aplicação de fertilizantes para atingir melhorias na produtividade no milho safrinha representa uma elevada parcela do custo variável de produção, chegando a 30% no Estado de Mato Grosso (IMEA, 2016) e até 38% em São Paulo (PECCHIO, 2015).

Por meio do melhoramento genético, é possível desenvolver cultivares de milho mais eficientes no uso de nitrogênio (EUN), ou seja, que produzam mais ou a mesma quantidade de grãos, com menor demanda por N. O aumento da EUN não implica apenas benefícios econômicos, mas também reduz a poluição ambiental. As perdas de N por lixiviação de nitrato chegam a 6% (PEDROSA, 2013), enquanto que as por volatilização podem ultrapassar 50% do total aplicado na cultura (TASCA et al., 2011).

Trabalhos com melhoramento para a EUN já vêm sendo realizados, simultaneamente, em condições ótimas e estressantes de cultivo, através da utilização de índices que reúnem o desempenho dos genótipos nos dois ambientes. Um importante índice

nesse contexto é o Índice de Tolerância à Baixa Disponibilidade de Nitrogênio (ITBN), proposto por Miti et al. (2010). Este relaciona a produtividade de grãos nas duas condições (disponibilidade ideal e baixa), obtendo o percentual de redução da produtividade do ambiente estressante em relação ao ambiente ótimo. Este índice vai ao encontro do conceito de tolerância ao estresse, sendo que, quanto menor a perda de produtividade da condição estressante para a ótima, menor será o valor do índice e maior será a tolerância da planta ao estresse.

Neste contexto, caracteres relacionados à morfologia do sistema radicular podem fornecer informações relevantes sobre a eficiência no uso do nitrogênio em genótipos de milho (MAJEROWICZ et al., 2002). Isto porque o sistema radicular exerce papel essencial na captação do nitrogênio via água (HORST et al., 2003; MACKAY; BARBER, 1986; SATTELMACHER et al., 1990). Assim, sob estresse nutricional ou hídrico, por exemplo, um maior desenvolvimento radicular pode auxiliar a obter o nutriente em camadas de solo mais profundas e, dessa forma, evitar a deficiência de nitrogênio (CHUN et al., 2005). Entretanto, a influência genética de caracteres de raiz na eficiência no uso do nitrogênio ainda é pouco conhecida, tanto a nível fenotípico como genômico.

Além destes, caracteres fisiológicos também podem exercer contribuição importante para a seleção de genótipos eficientes no uso de nitrogênio. A análise de taxa fotossintética, por exemplo, pode ser um indicador de uma planta com alta ou baixa eficiência na absorção e utilização de N, dada a participação direta do nitrogênio na síntese de Rubisco e clorofila (EPSTEIN; BLOOM, 2006).

A avaliação de genótipos em relação à EUN através desses caracteres por métodos tradicionais pode ser muito informativa, mas se baseia em análise destrutiva dos órgãos vegetais envolvidos na maioria das vezes, o que implica alta demanda de tempo e custo, além de alta contribuição de efeitos ambientais nos resultados, indesejados em programas de melhoramento (VARSHNEY et al., 2011). Esses desafios tornam a seleção precoce direta uma alternativa atrativa para essas características morfofisiológicas, pois permite a realização de seleção de forma rápida, com menor custo e com maximização do ganho de seleção (FRITSCHÉ-NETO; BORÉM, 2011).

Assim, em conjunto com uma caracterização morfofisiológica, a caracterização detalhada dos genótipos com base em dados moleculares pode contribuir estrategicamente para otimizar a utilização do germoplasma disponível em programas de melhoramento de milho (ROMAY et al., 2013), assim como para o entendimento da base genética de caracteres de interesse. Para isso, os estudos de associação genômica ampla (GWAS – *Genome-wide*

Association Studies) representam um método de alta resolução para mapear *loci* quantitativos com base em desequilíbrio de ligação (LD) (DHANAPAL et al., 2015). Essa abordagem estima a associação entre locos e o caráter fenotípico em nível populacional por meio de testes de hipóteses visando detectar efeitos de SNPs (*Single Nucleotide Polymorphism*) significativos.

Em milho, dado o rápido decaimento do LD (aproximadamente 2 kb) e a grande diversidade genética, o GWAS evidencia-se como um método acurado para identificação de genes específicos. Prova disso é que a metodologia GWAS vem sendo extensivamente usada em estudos genéticos de plantas para condições de estresses abióticos (BOUCHET et al., 2013; XUE et al., 2013; ZHANG et al., 2013) e características agronomicamente importantes do milho, como arquitetura foliar (TIAN et al., 2011), concentração e composição do óleo da semente (LI et al., 2013), resistência a doenças (GOWDA et al., 2015; MAMMADOV et al., 2015), desenvolvimento radicular (PACE et al., 2015) e altura de planta (PEIFFER et al., 2014).

1.1. Objetivos

1.1.1. Objetivo geral

Identificar regiões do genoma do milho tropical associadas a caracteres morfológicos de raiz, a parâmetros fisiológicos do fotossistema II e ao índice de resposta da planta ao estresse por nitrogênio.

1.1.2. Objetivos específicos

- Identificar marcadores SNP associados a índice de resposta da planta ao estresse por nitrogênio e a caracteres de morfologia de raiz e fisiologia do fotossistema II;
- Fazer a anotação funcional dos SNP identificados como significativos e identificar possíveis fatores moleculares que controlam estas características.
- Comparar a coincidência e/ou localização das marcas significativas nos diferentes caracteres no genoma, assim como possíveis relações genômicas entre elas.

2. REVISÃO BIBLIOGRÁFICA

2.1. Cenário nacional do cultivo de milho

Atualmente, o cultivo de milho em território nacional abrange 16,2 milhões de hectares com produção de 84,6 milhões de toneladas de grãos (CONAB, 2016), tornando o Brasil o terceiro maior produtor mundial da cultura. O país diferencia-se por produzir duas safras ao ano sem o uso da irrigação, sendo a segunda safra superior à primeira desde 2012.

O milho de segunda safra, também conhecido como milho safrinha, é definido como aquele semeado entre os meses de janeiro e março. Esta modalidade de cultivo começou no estado do Paraná no final da década de 1970, quando se buscava uma alternativa ao cultivo de cereais de inverno após a colheita da soja (DUARTE, 2008). No início dos anos 2000, o plantio de milho segunda safra teve um grande crescimento, decorrente principalmente do plantio no estado do Mato Grosso, feito por produtores oriundos do Sul do País.

O advento e a tecnificação do milho safrinha foram responsáveis pela transformação da cultura de milho no Brasil. Houve mudança espacial, com o avanço da cultura para o Centro-Oeste e recentemente para os chapadões do Maranhão, Piauí e Tocantins, e temporal, com a maior parte da área de milho verão sendo substituída pela de soja, passando a ser cultivado preferencialmente na segunda safra (DUARTE; KAPPES, 2015).

Apesar da consolidação do milho de segunda safra na produção nacional, seu cultivo ainda apresenta riscos, principalmente em relação a diversos tipos de estresses abióticos. Destaca-se entre eles a deficiência de nitrogênio (N). Esse estresse nutricional afeta negativamente o desenvolvimento das plantas, reduzindo a produtividade e condicionando uma alta aplicação de adubação nitrogenada (FRITSCHÉ-NETO; BORÉM, 2012). No país, a cultura do milho é a principal consumidora desse adubo, com utilização em torno de 1,28 milhões de toneladas por ano (PETROBRAS, 2009). Assim, o desenvolvimento de cultivares eficientes no uso de N apresenta-se como uma opção economicamente viável e ecologicamente sustentável, possibilitando produtividade satisfatória em condições de estresse nutricional e, portanto, com menor custo de produção e menor impacto ambiental (ABDEL-GHANI et al., 2013; XU et al., 2014).

2.2. Melhoramento com foco em eficiência no uso de nitrogênio

Por participar de processos biológicos essenciais das plantas, a deficiência de N geralmente provoca redução da área foliar, da fotossíntese e da biomassa final e,

consequentemente, na produtividade de grãos (BOOMSMA et al., 2009; CIAMPITTI; VYN, 2011). Assim, há uma enorme demanda por alternativas sustentáveis, como a utilização de plantas mais eficientes no uso de N (EUN), ou seja, que apresentem produtividades satisfatórias em baixa disponibilidade desse nutriente, ou ainda, que produzam a mesma quantidade, mas com menor demanda de adubo.

Os cultivares modernos de milho foram desenvolvidos em condições de alta fertilidade de solo, com foco em alta produtividade. O germoplasma de milho selecionado em condições ótimas de cultivo geralmente não é adequado para utilização em condições de estresses abióticos como a deficiência nutricional de nitrogênio (CECARELLI, 1996; FIDELIS et al., 2007). Entretanto, alguns autores têm constatado ampla diversidade genética entre plantas quanto a mecanismos fisiológicos e morfológicos da raiz e de parte aérea, quando elas são submetidas às condições adversas de fertilidade do solo (DOVALE et al., 2012; MEDICI et al., 2004).

Alguns genes importantes que controlam a produtividade de grãos em condições de estresses abióticos são diferentes daqueles em condições ótimas (ATLIN; FREY, 2000; FRITSCHÉ-NETO et al., 2010; SOUZA et al., 2008), o que faz com que um material selecionado em condições favoráveis não necessariamente seja o melhor em condições de estresse (SOUZA et al., 2009). Portanto, para que se obtenha maior sucesso na seleção para EUN, a mesma deve ser feita submetendo as plantas não apenas às condições ideais, mas também, à condição estressante. Nesse contexto, o uso de métodos de seleção precoce e seleção indireta para EUN é de grande interesse, pois aceleram o processo de seleção, descartando de imediato os genótipos inferiores e possibilitam concentrar os recursos nos potencialmente superiores (FRITSCHÉ-NETO et al., 2012).

Dentre as muitas características que tem potencial para serem usadas na seleção para EUN, pode-se destacar o maior desenvolvimento do sistema radicular. Esse caráter é altamente responsivo à disponibilidade de nutrientes e sua distribuição no solo (LINKOHR et al., 2002). Estudos demonstram que a extensão e natureza da variação genética do sistema radicular possuem grande implicação na obtenção e absorção de água e nutrientes para o aumento de produtividade sob estresse hídrico ou baixa fertilidade do solo (TOALDO et al., 2013). Assim, o crescimento espacial do sistema radicular determina a habilidade da planta em explorar recursos que estão mal distribuídos. Com isso, o comprimento radicular pode ser um indicador do potencial de absorção de água e nutrientes, sendo proporcionalmente maior o volume ocupado e explorado do solo quanto maior for o comprimento deste (ZONTA et al., 2006).

A fotossíntese também pode ser um indicador de uma planta com alta ou baixa eficiência na absorção e utilização de N, atuando como um caráter fisiológico que pode discriminar genótipos quanto a EUN. Sua inferência pode ocorrer através de dados de fluorescência de clorofila, visto que o N é um elemento importante para a síntese da proteína Rubisco e da clorofila (EPSTEIN; BLOOM, 2006). Logo, em condições de baixa disponibilidade de N, ocorre a redução da taxa fotossintética (BOOMSMA et al., 2009; CIAMPITTI; VYN, 2011), e consequentemente, pode haver redução do período de crescimento vegetativo e no índice de colheita (SINCLAIR; GARDNER, 1998).

Outra abordagem é a utilização de índices de resposta de plantas ao estresse, expressões matemáticas que compõem um valor que auxilia na seleção dos melhores genótipos para uma determinada característica. Com estes, a seleção pode ser feita observando o desempenho das plantas na condição favorável e em estresse, usando os *scores* que agrupam ambas as situações em apenas um valor (WU et al., 2011). O Índice de Tolerância à Baixa Disponibilidade de Nitrogênio (ITBN) é um exemplo, o qual relaciona a produtividade de grãos nas duas condições, obtendo o percentual de redução da produtividade do ambiente estressante em relação ao ambiente ótimo (MITI et al., 2010). Este índice vai ao encontro do conceito de tolerância ao estresse, sendo que, quanto menor a perda de produtividade da condição estressante para a ótima, menor será o valor do índice e mais tolerante a planta será ao estresse.

2.3. Estudos de Associação Genômica Ampla (GWAS)

Além de uma acurada caracterização fenotípica, a caracterização detalhada dos genótipos com base em dados moleculares pode contribuir estrategicamente para otimizar a utilização do germoplasma disponível em programas de melhoramento de milho visando alta EUN (ROMAY et al., 2013). Por isso, visando o aperfeiçoamento da relação custo-benefício, a otimização do tempo e a realização de uma seleção com acurácia, deve-se lançar mão de metodologias precisas que amostram o genoma e o correlacionem com as características fenotípicas.

Contextualizam-se nesse cenário os marcadores moleculares, cuja disponibilidade e custo cada vez mais acessível possibilitam seu amplo emprego nos programas de melhoramento, excluindo a necessidade de avaliações fenotípicas contínuas (BERNARDO; YU, 2007). Dentre eles, destacam-se os de Polimorfismo de um Único Nucleotídeo (*Single-Nucleotide Polymorphism* - SNP). Os SNP baseiam-se na detecção de polimorfismos

resultantes da alteração de uma única base no genoma e têm sido amplamente utilizados, pois apresentam diversas vantagens em relação a outros tipos: é a forma de variação mais abundante no genoma, apresentam alta reprodutibilidade, possibilidade de automação, ampla cobertura do genoma e baixo custo. Assim, os SNP vêm sendo empregados com sucesso em estudos de diversidade e estrutura genética de populações, construção de mapas genéticos (GORE et al., 2009), detecção de genes/QTLs e análise de arquitetura genética de características complexas (COOK et al., 2012; PEIFFER et al., 2014).

Dada a relevância de unir análises fenotípicas e fisiológicas a moleculares, diferentes abordagens têm sido utilizadas. De forma geral, elas implicam genotipagem com alta densidade de marcadores no genoma e o uso de métodos baseados em desequilíbrio de ligação (LD). Com o rápido desenvolvimento das tecnologias de sequenciamento e métodos computacionais, emerge a GWAS (*Genome Wide Association Studies*), utilizada como uma poderosa ferramenta para detectar a variação natural relacionada aos caracteres complexos em diversas culturas (RAFALSKI, 2010).

Em populações genéticas, LD é a associação não aleatória de alelos em *loci* diferentes. Sua extensão determina não apenas a resolução de mapeamento que pode ser alcançada, mas também o número de marcadores necessários para uma cobertura adequada do genoma em um estudo de associação genômica (WHITT; BUCKLER, 2003).

A abordagem GWAS permite a identificação de regiões cromossômicas associadas a um determinado fenótipo sem conhecimento da mutação causal da variação genética, de forma que os SNPs identificados podem não ter efeito direto sobre a característica em estudo (KORTE; FARLOW, 2013). Como não são feitas suposições sobre a localização genômica das variantes causais, essa abordagem estima a associação entre locos e o caractere fenotípico em nível populacional por meio de testes de hipóteses, visando detectar efeitos com significância estatística (RESENDE et al., 2013).

Nesse contexto, os estudos GWAS para identificação de QTLs podem auxiliar o desenvolvimento mais rápido de cultivares com maior EUN, pois permitem o entendimento do controle genético e fisiológico dessa característica. Isso porque, após a identificação dos SNPs em LD com caracteres relacionados a EUN, é possível o monitoramento de regiões cromossômicas específicas que afetam a expressão do caráter, fornecendo informações sobre o número e a localização dos locos que o controlam (FRITSCHÉ-NETO; BORÉM, 2011). Ainda, por meio do uso de bancos de dados, é possível realizar a caracterização dos marcadores significativos resultantes através de anotação funcional.

3. MATERIAL E MÉTODOS

3.1. Multiplicação das linhagens

Para se obter sementes em quantidade suficiente e com alta qualidade para a implantação dos experimentos, primeiramente, realizou-se a multiplicação de sementes de sessenta e quatro linhagens de milho tropical. Para isso, as parcelas foram constituídas por linhas de quatro metros de comprimento, com espaçamento de 0,85 metros entre linhas e 0,20 metros entre plantas. Foram semeadas 20 sementes por linha. O manejo e tratos culturais foram realizados conforme recomendado para a cultura do milho (EMBRAPA, 2016). Esta etapa foi realizada na área experimental do Departamento de Genética da ESALQ/USP, em Piracicaba – SP (22°42'23"S, 47°38'14"W, 535m).

As linhagens são oriundas da Universidade Federal de Viçosa (UFV), onde foram desenvolvidas por meio de seleção disruptiva para eficiência no uso de nitrogênio (EUN), de modo que apresentam grande variabilidade genética para a característica em questão (SANT'ANA, 2015).

3.2. Instalação e condução dos experimentos no campo

As sessenta e quatro linhagens (geração S7-S9) foram avaliadas na Estação Experimental de Anhumas (22°50'51.96"S, 48°01'06"W, 466m), pertencente a ESALQ/USP em Anhembi – SP, na segunda safra de 2014 e 2015, e no campo experimental do Departamento de Genética (22°42'28"S, 47°38'11"W, 539m), da ESALQ/USP em Piracicaba – SP, na segunda safra de 2015. A Estação Experimental de Anhumas corresponde a um ambiente típico dos cerrados brasileiros, com topografia plana e solo do tipo Latossolo Amarelo Distrófico. Já o campo experimental do Departamento de Genética possui solo do tipo Nitossolo Vermelho Eutroférico típico (NVEf1), textura argilosa, com horizonte A moderado (análise de solo disponível em Anexo A).

Considerou-se cada combinação ano-local com sendo um ambiente, totalizando, portanto, três ambientes distintos (Anhembi 2014, Anhembi 2015 e Piracicaba 2015). Os dados meteorológicos do Ambiente 3 encontram-se no Anexo B.

O delineamento experimental foi o látice simples (8×8), com duas unidades experimentais em cada local: uma cultivada com disponibilidade ideal de nitrogênio (IN) e outra cultivada com baixa disponibilidade de nitrogênio (BN). Cada unidade foi composta por duas repetições. As parcelas foram constituídas de uma linha de quatro metros de

comprimento na safrinha 2014 e cinco metros na safrinha 2015, diferença adotada devido à necessidade de um maior número de plantas para avaliação do comprimento radicular. Em ambos os anos, utilizou-se espaçamento de 0,85 metros entre linhas e 0,20 metros entre plantas, havendo portanto 20 plantas por linha na safrinha 2014 e 25 na safrinha 2015.

A adubação de plantio foi de 300 kg ha⁻¹ do formulado NPK 4-14-8. Para determinar as doses de nitrogênio (N) adequadas para cada disponibilidade, levou-se em consideração o efeito do nitrogênio para redução de, em média, 50% da produtividade do milho (DOVALE et al., 2011; GALLAIS; HIREL, 2004). As doses de adubação foram diferentes de um ano para outro, mas igualmente proporcionais. No ano de 2014, foram aplicados 60 kg ha⁻¹ de N para a unidade experimental de baixa disponibilidade (sendo 27 kg ha⁻¹ aplicados na semeadura e 33 kg ha⁻¹ em cobertura) e 150 kg ha⁻¹ de N para a unidade experimental de disponibilidade ideal (27 kg ha⁻¹ aplicados na semeadura e 133 kg ha⁻¹ em cobertura). Já em 2015, esses valores corresponderam a 35 kg ha⁻¹ de N (sendo 12 kg ha⁻¹ aplicados na semeadura e o 23 kg ha⁻¹ em cobertura) e 100 kg ha⁻¹ de N (12 kg ha⁻¹ aplicados na semeadura e o restante em cobertura), respectivamente. Tanto em IN quanto em BN, foram feitas duas adubações de cobertura: aos 30 dias (estádio V6) e aos 50 dias (estádio VT) após a semeadura. Para assegurar que o estresse nas plantas ocorresse apenas em decorrência da deficiência do nitrogênio e que o erro experimental fosse minimizado, todos os outros fatores foram controlados por meio de irrigação, tratamentos fitossanitários, adubação e manejo adequados, de acordo com os procedimentos recomendados para a cultura do milho (EMBRAPA, 2016).

3.3. Fenotipagem

Em cada unidade experimental, foram coletados os dados de estande (número de plantas na parcela no estágio de maturação fisiológica). Em seguida, com a realização da colheita, foram coletados os dados de produtividade de grãos (PG) em kg ha⁻¹ e de umidade de grãos (%). Por meio desses dados, a produtividade foi corrigida para umidade de grãos com base seca a 13%. A correção da produtividade com base no estande não foi realizada, pois este último apresentou efeito genético significativo.

No ano de 2015, adicionalmente à PG, foram também coletados dados morfofisiológicos dos caracteres descritos a seguir:

- Comprimento de raiz: considerou-se o comprimento total de raiz (CRa), correspondente à soma dos comprimentos de todas as raízes de cada planta, avaliado no estágio V8 pela amostragem de duas plantas por parcela. A mensuração, em centímetros, foi

realizada através do software WinRHIZO™ Arabidopsis acoplado a um scanner EPSON LA2400, com definição de 2400 dpi.

- Fluorescência de clorofila: foi avaliada a eficiência fotossintética do fotossistema II (EFPSII), estimada por meio do fluorômetro portátil compacto para medição de fotossíntese JUNIOR-PAN. As informações coletadas foram obtidas entre às 9 e 12h da manhã, em uma folha por planta (terceira folha expandida) em duas plantas aleatórias dentro da parcela. Seguindo a metodologia descrita por DaMatta et al. (1997), as folhas permaneceram no escuro durante 30 minutos antes da coleta de dados. Para isso, foram empregados grampos com áreas cilíndricas opacas que mantiveram a área mensurada sob completa ausência de luz. Após os centros de reações estarem completamente abertos, determinou-se a EFPSII por:

$$EFPSII_{\phi} = \frac{F_V}{F_M}$$

em que $F_V = F_M - F_0$ é a fluorescência variável, sendo F_M a fluorescência máxima e F_0 a fluorescência inicial.

3.4. Cálculo do Índice de Tolerância à Baixa disponibilidade de Nitrogênio

A partir dos dados de PG em cada unidade experimental, foi estimado o Índice de Tolerância à Baixa disponibilidade de Nitrogênio (ITBN) para cada genótipo em cada repetição, descrito por Miti et al. (2010):

$$ITBN_{ij} = \left(1 - \frac{PG_{(BN)ij}}{PG_{(IN)ij}} \right) \times 100$$

em que $ITBN_{ij}$ é o índice de tolerância à baixa disponibilidade de nitrogênio da linhagem i na repetição j ; $PG_{(BN)ij}$ é a produtividade de grãos no ambiente de baixa disponibilidade de nitrogênio (valor fenotípico) da linhagem i na repetição j ; $PG_{(IN)ij}$ é a produtividade de grãos no ambiente de disponibilidade ideal de nitrogênio (valor fenotípico) da linhagem i na repetição j .

3.5. Análises genético-estatísticas dos dados fenotípicos

As análises estatísticas dos dados fenotípicos foram realizadas considerando-se equações de modelos mistos. Para isso, utilizou-se o método da Máxima Verossimilhança Restrita/Melhor Preditor Linear Não-viesado (REML/BLUP) para a estimação dos componentes de variância e predição dos efeitos aleatórios. A significância dos efeitos fixos foi feita pelo teste de Wald, enquanto a dos efeitos aleatórios foi feita pelo teste da razão de verossimilhança (*Likelihood Ratio Test* – LRT), por meio da Análise de Deviance (ANADEV), obtida através do pacote ASReml-R® (GILMOUR et al., 2009) do software R (R CORE TEAM, 2014).

Primeiramente, realizou-se a análise individual para cada disponibilidade de nitrogênio. Os componentes de variância e os valores genotípicos foram obtidos considerando-se o seguinte modelo:

$$\mathbf{y} = \mathbf{X}\mathbf{r} + \mathbf{Z}\mathbf{g} + \mathbf{W}\mathbf{b} + \mathbf{T}\mathbf{i} + \boldsymbol{\varepsilon}$$

em que $\mathbf{y}$ é o vetor de BLUP das linhagens para as variáveis; $\mathbf{r}$ é o vetor do efeito de ambiente e repetição dentro de ambiente, somados à média geral, considerado como fixo; $\mathbf{g}$ é o vetor do efeito de linhagem, considerado como aleatório, em que $\mathbf{g} \sim N(0, I\sigma_g^2)$; $\mathbf{b}$ é o vetor do efeito de bloco dentro de repetição, considerado como aleatório, em que $\mathbf{b} \sim N(0, I\sigma_b^2)$; $\mathbf{i}$ é o vetor do efeito da interação linhagem $\times$ ambiente, considerado como aleatório, em que $\mathbf{i} \sim N(0, I\sigma_i^2)$; $\boldsymbol{\varepsilon}$ é o erro experimental, em que $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$. $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{W}$ e $\mathbf{T}$ são matrizes de incidência que relacionam os efeitos dos vetores independentes de cada matriz ao vetor dependente $\mathbf{y}$.

Em seguida, foi realizada a análise conjunta com o fator “nível de nitrogênio” correspondente às disponibilidades ideal (IN) e baixa (BN) de nitrogênio. Os componentes de variância e os valores genotípicos foram obtidos considerando-se o modelo acima descrito com inclusão dos vetores $\mathbf{n}$ (vetor do efeito de nível de nitrogênio) e $\mathbf{x}$ (vetor do efeito da interação linhagem $\times$ nível de nitrogênio), pelo qual tem-se:

$$\mathbf{y} = \mathbf{X}\mathbf{r} + \mathbf{L}\mathbf{n} + \mathbf{Z}\mathbf{g} + \mathbf{W}\mathbf{b} + \mathbf{T}\mathbf{i} + \mathbf{H}\mathbf{x} + \boldsymbol{\varepsilon}$$

em que $\mathbf{y}$ é o vetor de BLUP das linhagens para as variáveis; $\mathbf{r}$ é o vetor do efeito de ambiente e repetição dentro de ambiente, somados à média geral, considerado como fixo; $\mathbf{n}$ é o vetor do

efeito de nível de nitrogênio, considerado como fixo; $\mathbf{g}$ é o vetor do efeito de linhagem, considerado como aleatório, em que $\mathbf{g} \sim N(0, I\sigma_g^2)$; $\mathbf{b}$ é o vetor do efeito de bloco dentro de repetição, considerado como aleatório, em que $\mathbf{b} \sim N(0, I\sigma_b^2)$; $\mathbf{i}$ é o vetor do efeito da interação linhagem $\times$ ambiente, considerado como aleatório, em que $\mathbf{i} \sim N(0, I\sigma_i^2)$; $\mathbf{x}$ é vetor do efeito da interação linhagem $\times$ nível de nitrogênio, considerado como aleatório, em que $\mathbf{x} \sim N(0, I\sigma_x^2)$; $\boldsymbol{\varepsilon}$ é o erro experimental, em que $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$. X , L , Z , W , T e H são matrizes de incidência que relacionam os efeitos dos vetores independentes de cada matriz ao vetor dependente $\mathbf{y}$.

A herdabilidade no sentido amplo de cada variável foi estimada com base nos componentes de variância, por meio de:

$$h_g^2 = \frac{\sigma_g^2}{\left(\sigma_g^2 + \frac{\sigma_{g \times a}^2}{l} + \frac{\sigma_\varepsilon^2}{rl}\right)}$$

em que σ_g^2 é a variância genética, $\sigma_{g \times a}^2$ é a variância da interação genótipo $\times$ local, σ_ε^2 é a variância residual, r é o número de repetições e l é o número de locais.

A acurácia experimental foi estimada segundo Resende e Duarte (2007) por:

$$\hat{r} = \left[1 - \frac{1}{1 + rl \cdot CV_R^2}\right]^{1/2}$$

em que $CV_R = \frac{CV_g}{CV_\varepsilon}$ é o coeficiente de variação relativa, sendo o coeficiente de variação genotípica $CV_g = (\sqrt{\sigma_g^2}/\mu) \times 100$ e o coeficiente de variação residual $CV_\varepsilon = (\sqrt{\sigma_\varepsilon^2}/\mu) \times 100$; μ é a média geral da característica; σ_g^2 é a variância genética; σ_ε^2 é a variância residual; r é o número de repetições e l é o número de locais.

A partir dos valores genotípicos (BLUPS) das linhagens para as variáveis consideradas, foram obtidas correlações de Pearson, por meio da equação:

$$r = \frac{\sum(x_i) \times (y_i)}{\sqrt{\sum(x_i)^2 \times \sum(y_i)^2}}$$

em que r é o coeficiente de correlação de Pearson e x_i e y_i são os valores genotípicos das linhagens para as variáveis X e Y, respectivamente.

3.6. Genotipagem e controle de qualidade dos dados genômicos

A coleta do material genético foi realizada no estádio de três folhas completamente expandidas (V3). O DNA genômico das linhagens foi coletado e isolado utilizando-se o protocolo adaptado de Murray e Thompson (1980). As amostras foram genotipadas com a plataforma *Affymetrix® Axiom® Maize Genotyping Array* (UNTERSEER et al., 2014) com 616.201 marcadores do tipo SNP.

O conjunto de marcadores obtidos na genotipagem passou por uma fase de controle de qualidade baseada nos procedimentos de *Call Rate* (CR) e *Minor Allele Frequency* (MAF). Foram eliminados marcadores com mais de 5% de valores perdidos ($CR < 0,95$), locos pouco variáveis na população considerando a frequência do alelo menos frequente ($MAF < 0,05$), e marcadores com pelo menos um loco heterozigoto, utilizando-se o pacote HapEstXXR-R (KNUEPPEL et al., 2015).

3.7. Análises de Associação Genômica Ampla (GWAS)

As análises GWAS baseadas em SNP foram realizadas utilizando-se modelo linear misto, nos quais considerou-se tanto a estrutura populacional (matriz Q, obtida por meio de componentes principais), quanto a relação de parentesco entre as linhagens (matriz K), visando minimizar a ocorrência de associações espúrias:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \boldsymbol{\varepsilon}$$

em que $\mathbf{y}$ é o vetor de valores genotípicos (BLUPS) das linhagens para os caracteres de interesse; $\boldsymbol{\beta}$ é o vetor contendo os efeitos fixos, incluindo a média geral de efeito dos marcadores genéticos e a estrutura populacional (Q); $\mathbf{v}$ é o vetor aleatório dos efeitos genéticos aditivos ponderados pela matriz de parentesco (K); $\boldsymbol{\varepsilon}$ é o vetor residual. Assume-se que os vetores $\mathbf{v}$ e $\boldsymbol{\varepsilon}$ seguem distribuição normal, média zero e variância dada por:

$$Var\begin{pmatrix} \mathbf{v} \\ \boldsymbol{\varepsilon} \end{pmatrix} = \begin{pmatrix} G & 0 \\ 0 & R \end{pmatrix}$$

em que $G = \sigma_a^2 K$, sendo σ_a^2 a variância aditiva; $R = I\sigma_\varepsilon^2$, sendo σ_ε^2 a variância residual, assumida como homogênea. O número de componentes principais usados para compor a matriz Q variou conforme a característica, sendo na maioria dos casos igual a 2.

Para definir o p-valor de cada SNP e sua significância, assumiu-se o *threshold* por correção de Bonferroni, em que $p \leq 0,01/n$, sendo n = número total de SNP.

Todas as etapas de GWAS foram conduzidas no software R, por meio da utilização dos pacotes GAPIT (WANG et al., 2014) e FarmCPU (LIU, X. et al., 2016).

3.8. Desequilíbrio de ligação (LD)

O desequilíbrio de ligação (LD) foi calculado par a par entre dois SNP utilizando-se uma janela de 50 kb. Foram utilizadas as duas principais medidas para o cálculo desse evento:

- D' , coeficiente padronizado que mensura a história recombinação, útil para acessar a probabilidade de recombinação histórica numa dada população, conforme segue:

$$D' = \begin{cases} \frac{D}{\min(p_A \times p_b, p_a \times p_B)}, D > 0 \\ \frac{D}{\min(p_A \times p_B, p_a \times p_b)}, D < 0 \end{cases}$$

em que, $D = P_{AB} - p_A \times p_B$, sendo p_A, p_a, p_B e p_b as frequências alélicas de A, a, B e b , respectivamente, e P_{AB} a frequência do haplótipo AB .

- r^2 , que representa o quadrado do coeficiente da correlação entre os alelos de dois loci, útil no contexto de estudos de associação e mensurado através da seguinte equação:

$$r^2 = \frac{(P_{AB} \times P_{ab} - P_{Ab} \times P_{aB})^2}{(p_A \times p_a \times p_B \times p_b)}$$

em que p_A, p_a, p_B e p_b são as frequências dos alelos A, a, B e b , respectivamente, e P_{AB}, P_{ab}, P_{Ab} e P_{aB} são as frequências dos haplótipos AB, ab, aB e Ab , na população, respectivamente.

As análises de LD foram executadas nos softwares TASSEL (BRADBURY et al., 2007) e R (R CORE TEAM, 2014).

3.9. Análise de genes candidatos

Para identificar e analisar genes candidatos, utilizou-se a terceira versão da montagem genômica da linhagem temperada B73, atualmente estabelecida como genoma de referência (SCHNABLE et al, 2009).

A abrangência da janela genômica a ser percorrida a partir da posição física de cada SNP significativo para a identificação de genes candidatos foi determinada em função do LD médio considerando-se todos os cromossomos (FARFAN et al., 2015; LIU, N. et al., 2016).

Após a delimitação da região a ser considerada, foram apontados os possíveis genes candidatos que flanqueiam o marcador em questão, assim como os tipos de metabólitos, rotas em que estão associados e seus potenciais efeitos e relações fisiológicas com a EUN. A anotação funcional dos genes candidatos foi realizada utilizando-se os bancos de dados MaizeGDB (<http://www.maizegdb.org>), NCBI (<https://www.ncbi.nlm.nih.gov>) e UniProt (<http://www.uniprot.org>).

4. RESULTADOS

4.1. Análise de Deviance e valores genotípicos

Após a colheita na safra de 2014, constatou-se a perda de um genótipo nos ensaios a campo. Assim, para o ITBN, o número final de materiais avaliados foi de sessenta e três. Para CRa e EFPSII, as linhagens foram avaliadas apenas no ano de 2015, mas considerando todos os sessenta e quatro genótipos.

Realizou-se primeiramente a análise de Deviance individual para as variáveis ITBN, CRa e EFPSII, considerando-se as duas últimas separadamente em IN e BN. Para ITBN e CRa em ambas as disponibilidades, as linhagens diferiram entre si de maneira significativa, diferentemente do que foi constatado para EFPSII (Tabela 1).

Tabela 1. Resumo da análise individual para as variáveis ITBN (Índice de Tolerância à Baixa disponibilidade de Nitrogênio), CRa (Comprimento de Raiz) e EFPSII (Eficiência Fotossintética do Fotossistema II) em disponibilidade ideal (IN) e baixa (BN) de nitrogênio em linhagens de milho tropical. Teste de Wald para os efeitos fixos e de LRT para os efeitos aleatórios, estimados via REML/BLUP

Fator de variação	ITBN	CRa IN	CRa BN	EFPSII IN	EFPSII BN
<i>Efeito Fixo</i>					
Ambiente	4,43 ^{ns}	37,95 ^{**}	16,65 ^{**}	6,20 [*]	0,8 ^{ns}
Repetição/Ambiente	1,09 ^{ns}	0,15 ^{ns}	11,91 ^{**}	3,60 ^{ns}	3,4 ^{ns}
<i>Efeito Aleatório</i>					
Bloco/Repetição	< 0,01 ^{ns}	< 0,01 ^{ns}	0,30 ^{ns}	8,50 [*]	0,59 ^{ns}
Linhagem	6,65 [*]	19,89 ^{**}	26,01 ^{**}	1,27 ^{ns}	3,37 ^{ns}
Linhagem x Ambiente	4,22 [*]	1,72 ^{ns}	1,80 ^{ns}	1,27 ^{ns}	< 0,01 ^{ns}

LRT – Likelihood ratio test; ns - não significativo; * - significativo a 5% pela LRT ou F Wald, ** - significativo a 1% pela LRT ou F Wald

Prosseguiu-se à análise conjunta para as variáveis CRa e EFPSII, considerando agora também o fator de variação “nível de nitrogênio” (Tabela 2). As linhagens diferiram significativamente em relação ao nível de nitrogênio para ambas as variáveis, indicando que as doses de N aplicadas para gerar ambientes contrastantes foram suficientes.

Novamente, não foram constatadas diferenças significativas para o efeito de linhagem em relação à EFPSII, indicando ausência de variabilidade genética no painel para este caráter (Tabela 2). Portanto, essa variável foi desconsiderada nas análises posteriores.

Tabela 2. Resumo da análise conjunta para as variáveis CRa (Comprimento de Raiz) e EFPSII (Eficiência Fotossintética do Fotossistema II) em linhagens de milho tropical. Teste de Wald para os efeitos fixos e teste de LRT para os efeitos aleatórios, estimados via REML/BLUP

Fator de variação	CRa	EFPSII
<i>Efeito Fixo</i>		
Ambiente	52,43**	4,90**
Nível N	7,29**	118,10**
Repetição/Ambiente	7,39**	4,60*
<i>Efeito Aleatório</i>		
Bloco/Repetição	< 0,01 ^{ns}	4,46*
Linhagem	83,30**	2,93 ^{ns}
Linhagem × Nível N	< 0,01 ^{ns}	1,16 ^{ns}
Linhagem × Ambiente	1,68 ^{ns}	1,21 ^{ns}

LRT – Likelihood ratio test; ns - não significativo; * - significativo a 5% pela LRT ou F Wald; ** - significativo a 1% pela LRT ou F Wald

Analisando-se os parâmetros genéticos, para CRa em BN foram obtidos os maiores valores de herdabilidade (0,54) e acurácia (0,78), além do menor coeficiente de variação (26,89%). O ITBN apresentou valores muito inferiores aos demais, com baixas herdabilidade e acurácia fenotípica, além de um alto coeficiente de variação, o que se justifica pelo fato do índice ser obtido por uma razão entre variáveis de produtividade (Tabela 3).

Tabela 3. Medidas de dispersão e parâmetros genéticos estimados em linhagens de milho tropical para Índice de Tolerância à Baixa disponibilidade de Nitrogênio (ITBN), comprimento de raiz em disponibilidade ideal (CRa_IN) e baixa (CRa_BN) de nitrogênio (em centímetros)

Parâmetros	ITBN	CRa_IN	CRa_BN
Média	26,26%	1434,80	1535,50
Desvio padrão	18,63%	501,18	520,27
Coeficiente de variação (CV%)	58,12	28,70	26,89
Herdabilidade (h_g^2)	0,13	0,49	0,54
Acurácia ($\hat{r}$)	0,47	0,75	0,78

Os coeficientes de correlação de Pearson entre valores genotípicos das linhagens para as variáveis ITBN, CRa em BN e CRa em IN indicaram correlação significativa positiva forte entre os CRa e positiva desprezível não significativa entre ITBN e CRa em ambas as disponibilidades (Tabela 4).

Tabela 4. Coeficientes de correlação entre valores genotípicos das variáveis consideradas: Índice de Tolerância à Baixa disponibilidade de Nitrogênio (ITBN), comprimento de raiz em disponibilidade ideal (CRa_IN) e baixa (CRa_BN) de nitrogênio

	ITBN	CRa_IN	CRa_BN
ITBN	-	0,11 ^{ns}	0,09 ^{ns}
CRa_IN		-	0,72 [*]
CRa_BN			-

ns - não significativo; * - significativo a 5% pelo teste T

4.2. Estrutura da população

A estrutura da população foi representada graficamente com base nos dois principais autovetores dos componentes principais da matriz de parentesco genômico (K). Para isso, primeiramente a matriz K foi calculada por meio da equação de VanRaden (2008). O componente principal 1 (PC1) explica 5,1% da variância total, enquanto o componente principal 2 (PC2) explica 4,6%. Os resultados indicaram ausência de subgrupos populacionais distintos (Figura 1).

Figura 1. Representação gráfica da estrutura populacional do painel de 64 linhagens de milho tropical. a) Análise de Componentes Principais (PCA) e b) *Heat Map* da análise de matriz de parentesco (K).

4.3. Desequilíbrio de ligação no painel de linhagens

O controle de qualidade realizado pelos procedimentos MAF, CR e eliminação de marcadores heterozigóticos reduziu o número de marcadores de 616.201 para 63.162, os quais foram utilizados nas análises genômicas e no cálculo do LD do presente painel. Todas as 64 linhagens consideradas apresentam alto grau de homozigose (média de 92,26%). A forte redução do número de marcadores (retenção de 9,75% do total) ocorreu porque qualquer SNP que apresentasse pelo menos um loco hetemórfico em uma das linhagens foi eliminado.

Um rápido declínio de LD foi observado com o aumento da distância física em todos os cromossomos e ao longo do genoma (Figura 2). Considerando-se todos eles, o valor médio observado foi de $r^2 = 0,13$, com um comprimento médio de decaimento LD permitindo a utilização de uma janela de abertura de 25 mil pares de base a partir da posição física do SNP significativo, *upstream* e *downstream*, para percorrer o genoma em busca de genes candidatos.

Figura 2. Decaimento do desequilíbrio de ligação (LD) em relação à distância (em mil bases). Cada ponto representa um par de distâncias entre dois marcadores na janela e seu coeficiente de correlação ao quadrado (r^2). A linha vermelha é a média móvel dos 10 marcadores adjacentes.

4.4. Associações dos SNP com as características fenotípicas

Visando minimizar correlações espúrias, foram consideradas no modelo estatístico para a análise de associação tanto a matriz Q, com informação de estrutura de população,

quanto a matriz K, com informações de parentesco entre as linhagens. Dessa forma, foi possível controlar falsos-positivos eficientemente, resultando em um modelo rigoroso evidenciado pela pequena quantidade de associações significativas (Figura 3).

Sete associações marcador-característica significativas foram observadas ($p \leq 1,58 \times 10^{-7}$) considerando-se todas as variáveis. Para ITBN, o número de marcas significativas foi quatro, uma no cromossomo 10, duas no cromossomo 7 e outra no 5. Para CRa em IN, dois SNP significativos foram detectados, um no cromossomo 4 e outro no 8. Para CRa em BN, detectou-se uma marca no cromossomo 7 (Figura 3).

Figura 3. *Manhattan plots* e *QQ-plots* resultantes de análise GWAS para Índice de Tolerância à Baixa disponibilidade de Nitrogênio (ITBN), comprimento de raiz em disponibilidade ideal (CRa_IN) e baixa (CRa_BN) de nitrogênio. *Manhattan plots* para ITBN, CRa_IN e CRa_BN são mostrados em *a*, *b* e *c*, respectivamente. *QQ plots* para ITBN, CRa_IN e CRa_BN são mostrados em *d*, *e* e *f*, respectivamente. A linha horizontal verde indica o limite de significância ($p = 1,58 \times 10^{-7}$).

Ao verificar uma associação significativa entre marcador e fenótipo, a plotagem dos *Manhattan plots* por meio do pacote FarmCPU indica apenas um marcador SNP ao invés de um pico (típico de modelos lineares gerais ou mistos), pois considera que as marcas adjacentes ao SNP mais significativo não apresentam significância. Portanto, o resultado gráfico é a apresentação de pontos ao invés de picos.

Com base na localização física dos SNP significativos na sequência genômica de referência B73, os genes candidatos concernentes aos *loci* significativos foram identificados (Tabelas 5 e 6).

Tabela 5. Gene candidato, identificação, posição física no cromossomo, p-valor, MAF e efeito de substituição alélica (ESA) dos SNPs associados significativamente com Índice de Tolerância à Baixa disponibilidade de Nitrogênio (ITBN), comprimento de raiz em disponibilidade ideal (CRa_IN) e baixa (CRa_BN) de nitrogênio

Variável	Gene candidato	SNP	Cr ^{*1}	Posição SNP	p-valor	MAF [§]	ESA ^{*2} (%)
ITBN	GRMZM2G136283	Affx.91269353	10	148504575	$2,71 \times 10^{-14}$	0,32	-8,0
	GRMZM2G131340	Affx.90760172	7	85581217	$4,01 \times 10^{-12}$	0,32	6,6
	GRMZM2G306935	Affx.90464448	7	98855426	$3,94 \times 10^{-9}$	0,22	-7,1
	GRMZM2G167431	Affx.91405933	5	200437599	$3,05 \times 10^{-8}$	0,21	4,6
CRa_IN	AC234185.1_FG006	Affx.90849540	8	72174334	$5,24 \times 10^{-8}$	0,36	5,7
	GRMZM2G099408	Affx.90303914	4	6448544	$1,25 \times 10^{-7}$	0,20	7,2
CRa_BN	GRMZM2G144971	Affx.90796403	7	172363612	$1,01 \times 10^{-7}$	0,11	10,3

*1: Cromossomo; §: *Minor allele frequency*; *2: Efeito de Substituição Alélica

Tabela 6. Posição física e anotação funcional dos genes candidatos para Índice de Tolerância à Baixa disponibilidade de Nitrogênio (ITBN), comprimento de raiz em disponibilidade ideal (CRa_IN) e baixa (CRa_BN) de nitrogênio

Variável	Gene candidato	Cr*	Posição do gene	Anotação
ITBN	GRMZM2G136283	10	148527506 a 148530904	GMP Sintetase
	GRMZM2G131340	7	85604861 a 85606104	Regulação da transcrição
	GRMZM2G306935	7	98838373 a 98844956	Reparo de DNA
	GRMZM2G167431	5	200432813 a 200435231	3-ketodihydrosphingosine redutase
CRa IN	AC234185.1_FG006	8	72172257 a 72173003	Fator transcricional MADS-box
	GRMZM2G099408	4	6442913 a 6443786	Fator transcricional MADS-box
CRa BN	GRMZM2G144971	7	172381310 a 172395119	Proteína associada ao microtúbulo

*: Cromossomo

4.5. Anotação funcional

A anotação funcional resultou em sete genes candidatos, um para cada marcador (Tabela 6). A análise indicou que três genes estão relacionados à regulação da transcrição gênica (GRMZM2G099408 no cromossomo 4, GRMZM2G131340 no cromossomo 7 e AC234185.1_FG006 no cromossomo 8), um está relacionado à síntese de Guanosina Monofosfato (GRMZM2G136283 no cromossomo 10), um está relacionado ao reparo de DNA (GRMZM2G306935 no cromossomo 7), outro relaciona-se ao processo de biossíntese de esfingolípídios (GRMZM2G167431 no cromossomo 5) e o outro está relacionado à ligação iônica de zinco (GRMZM2G144971 no cromossomo 7).

5. DISCUSSÃO

A eficiência no uso de nitrogênio é um caráter complexo regulado por muitos genes (KANT et al., 2011), e ainda pouco se sabia sobre a arquitetura genética dessa característica em milho. O painel considerado nesse trabalho foi constituído de linhagens que apresentam diversidade genética para EUN.

A extensão limitada de LD demanda um número substancialmente maior de marcadores genéticos para assegurar um poder adequado da análise visando detectar ligação entre um marcador e um loco causal (Hall et al., 2010). Assim, o LD médio relativamente baixo no painel contraposto pelo alto número de marcadores utilizado, aliado aos valores de herdabilidade (h_g^2) das variáveis estudadas, indicou boa resolução para identificação de SNP relacionados eficiência no uso de nitrogênio.

5.1. Análise exploratória dos dados

Em BN, o comprimento total de raiz foi maior (média de 1535,5 cm) do que em disponibilidade ideal do nutriente (média de 1434,8 cm). Esse mecanismo foi encontrado repetidamente em trabalhos prévios (CHUN et al., 2005; EGHBALL; MARANVILLE, 1993; LINKOHR et al., 2002; WANG et al., 2004; MORAIS, 2016) e, considerando-se que esses reportam um aumento da relação raiz/parte aérea, o maior comprimento de raiz em baixa disponibilidade do nutriente pode representar uma resposta adaptativa que conserva o nitrogênio da planta enquanto maximiza a habilidade de adquirir mais nitrogênio do solo. Normalmente, a resposta vegetal à deficiência de nitrogênio ocorre pelo aumento do comprimento total da raiz e pela alteração da arquitetura radicular, por meio do aumento da elongação de raízes laterais e axiais individuais e da redução do número de raízes axiais (GAUDIN et al., 2011).

O crescimento radicular sob baixa disponibilidade de nitrogênio depende da assimilação total de carbono e da alocação de carbono da raiz à parte aérea (CHUN et al., 2005). Assim, como pode haver uma grande divergência genotípica na habilidade de assimilação de carbono sob baixa disponibilidade de nitrogênio, é possível que o crescimento radicular seja potencializado, inalterado ou mesmo suprimido sob estresse de nitrogênio entre diferentes genótipos (COSTA et al., 2002; MACKAY; BARBER, 1986).

5.2. Associação ampla do genoma

Em relação ao ITBN, os dois genes candidatos identificados no cromossomo 7 referem-se a sequências gênicas cujos produtos proteicos ainda não foram totalmente elucidados. Para GRMZM2G131340, o polipeptídio resultante de 221 aminoácidos é encontrado no núcleo e está ligado ao processo biológico de regulação transcricional, com funções moleculares de ligação ao DNA e receptor dependente de ligação. Para GRMZM2G306935, o produto de 607 aminoácidos está envolvido no processo biológico de reparo de DNA, com funções moleculares de ligação à fita simples de DNA, ligação de ATP e atividade catalítica.

Visto que os genes são regulados diferentemente de acordo com o a disponibilidade de nitrogênio para a planta e que a regulação implica todos as fases de desenvolvimento e estruturas vegetais, não se pode dizer exatamente onde ou em qual momento essas proteínas estão agindo, uma vez que não foram caracterizadas completamente. Entretanto, é possível dizer que elas estão relacionadas à quantidade de nitrogênio disponível para a planta, pois seus respectivos genes candidatos foram detectados pela análise da relação da produtividade entre disponibilidades baixa e ideal do nutriente.

No cromossomo 5, o gene candidato identificado corresponde à enzima 3-ketodihydrosphingosine redutase. Ela está envolvida no processo de biossíntese de esfingolipídios, classe de lipídios participante em diversas atividades celulares dada sua variedade estrutural (LYNCH; DUNN, 2004). Em plantas, os esfingolipídios representam mais de 10% do lipídio total e estão presentes principalmente em tecidos não fotossintetizantes (DUNN et al., 2004). Apesar de seu papel ainda não ser bem definido, estudos recentes indicam que eles desempenham funções críticas, como componentes estruturais de membranas, moléculas bioativas envolvidas em transdução de sinal e metabólitos mediadores de processos celulares como morte celular programada (CHAO et al., 2011; COURSOL et al., 2003; MICHAELSON et al., 2016; WORRALL et al., 2008). Ademais, foram relatados em respostas das plantas a estresses por ataque de patógeno, hipóxia, resfriamento e congelamento (SAUCEDO-GARCIA et al., 2011; STEPONKUS; LYNCH, 1989; UEMURA et al., 1995; XIE et al., 2015).

Dessa forma, visto que os estudos sobre os esfingolipídios em plantas começaram a ser desenvolvidos recentemente e que já foi observada relação com estresses abióticos, é provável que eles também atuem em condições de estresse nutricional por nitrogênio, possibilidade corroborada pela verificação de uma associação significativa entre esse loco e o

ITBN pela GWAS. A atuação desses lipídios em processos diretamente ligados à interação e comunicação celular acentua essa inferência.

O gene candidato identificado no cromossomo 10 corresponde à enzima Guanosina Monofosfato Sintetase (EC 6.3.5.2 – GMP Synthase), pertencente à família das ligases, com funções moleculares de atividade de hidrólise da glutamina, ligação de ATP e formação de ligações carbono-nitrogênio. Ela promove a transferência do nitrogênio da glutamina, hidrolisada em glutamato e amônia na reação de glutaminase, a uma grande variedade de substratos (DUBOIS et al., 2003).

Assim como a enzima Glutamina Sintetase (EC 6.3.1.2 – GS), igualmente envolvida na interconversão glutamato-glutamina (EISENBERG et al., 2000; GALLAIS; HIREL, 2004), pode-se inferir que a GMP Synthase está relacionada à assimilação e reciclagem do nitrogênio mineral. A taxa de assimilação de amônio derivado da redução de nitrato e/ou a reciclagem de nitrogênio orgânico é de grande importância para a eficiência no uso de nitrogênio da planta (CREN; HIREL, 1999; LEA; IRELAND, 1999). Dessa forma, é possível a seleção indireta de genótipos tolerantes através da identificação do respectivo marcador, dada a forte associação entre o SNP correspondente a GMP Synthase ao ITBN. Gallais e Hirel (2004) observaram coincidências de QTL para produção e peso de grãos com QTL para GS, o que leva à hipótese de que essas enzimas desempenham um importante papel na determinação da produção.

Para comprimento de raiz em disponibilidade ideal de nitrogênio (CR em IN), os SNP significativos apontam como genes candidatos dois loci relacionados à regulação da transcrição. No cromossomo 4, a proteína de 247 aminoácidos atuante no núcleo é descrita como potencial fator de transcrição da família MADS-box, com função molecular de ligação ao DNA. Estudos desenvolvidos por Stelpflug et al (2016) mostram que essa proteína é encontrada apenas em sementes, sem expressão em outros órgãos do milho. O SNP significativo no cromossomo 8 identificou igualmente um gene candidato codante de um potencial fator de transcrição da família MADS-box. Nesse cromossomo, a proteína codificada contém 248 aminoácidos e é expressa majoritariamente na raiz, mais precisamente no meristema radicular.

Os fatores de transcrição do tipo MADS-box, que tem essa denominação devido às quatro primeiras famílias isoladas e caracterizadas (MCM1, AGAMOUS, DEFICIENS e SRF-serum response factor), foram isolados em muitos organismos eucariotos (THEISSEN et al., 1996). Em plantas, eles estão envolvidos no controle de importantes aspectos como o desenvolvimento do gametófito feminino e masculino, o desenvolvimento de embrião e da

semente e o desenvolvimento de raízes e flores (ALVAREZ-BUYLLA, 2000; PNUELI, 1994; REICHMANN; MEYEROWITZ, 1997). A identificação desses genes como candidatos analisando-se justamente o comprimento de raiz é compatível com os resultados já encontrados e evidencia sua participação no processo de desenvolvimento dessa estrutura. Em IN, dada a quantidade ideal de nitrogênio, não é necessário o crescimento demasiado do sistema radicular e este fator de transcrição provavelmente age diretamente nesse controle, através da regulação da expressão gênica.

Em relação ao comprimento de raiz em baixa disponibilidade de nitrogênio (CRa em BN), foi identificado um único SNP significativo remetente ao gene GRMZM2G14497, cuja caracterização é de proteína associada ao microtúbulo com função molecular de interação seletiva e não covalente a íons de zinco associada ao domínio Zinc finger CW-type, conhecido como “Dedo de Zinco”. Sua estrutura primária é extensa, com 1593 aminoácidos, e é expressa em todas as estruturas da planta de milho. As proteínas com o domínio Dedo de Zinco atuam em diversos processos como transcrição gênica, tradução, organização do citoesqueleto e remodelagem de cromatina (LAITY et al., 2001). Aqueles do tipo CW apresentam resíduos conservados de cisteína e triptofano e desempenham função especificamente na ligação ao DNA e/ou promovendo interações proteína-proteína em processos eucarióticos complexos como o estado de metilação da cromatina.

Dessa forma, os genes candidatos relacionados ao comprimento de raiz não apresentam função diretamente relacionada ao metabolismo do nitrogênio ou à eficiência no uso deste nutriente, mas apontam agentes regulatórios que afetam o crescimento em diferentes órgãos, e consequentemente, na demanda e disponibilidade de nitrogênio.

5.3. Relações genômicas entre genes candidatos

Não foram observadas marcas significativas coincidentes entre as variáveis analisadas. Isso pode ser explicado pelos coeficientes de correlação de Pearson obtidos. O ITBN apresenta correlação desprezível em relação ao CRa em IN e BN, enquanto que estes dois últimos, apesar de considerarem um mesmo caráter, representam condições contrastantes, o que justifica a presença de marcadores exclusivos em cada uma das situações.

A localização genômica de alguns dos genes candidatos sobrepôs aquela de QTL conhecidos, como foi o caso dos SNP identificados nos cromossomos 5 e 10. Mapeando-se genes específicos envolvidos no metabolismo de N para identificar possíveis genes candidatos, Gallais e Hirel (2004) verificaram que dois dos locos da enzima Glutamina

Sintetase Citosólica (GS1) coincide com vários QTLs para produtividade, peso de grãos, atividade da GS em folha, atividade da enzima Nitrato Redutase (NR) e conteúdo foliar de nitrato, sugerindo a participação dos genes dessas regiões em etapas primordiais na EUN. Os SNP significativos detectados nos cromossomos 5 e 10 por GWAS situam-se nas mesmas regiões cromossômicas dos locos de GS1 reportados por estes autores, evidenciando a relação dos genes nessas posições com a eficiência na absorção e utilização do nutriente.

Os genes detectados apresentaram efeitos pequenos, típicos de característica complexa, controlando uma porção relativamente pequena da variação fenotípica. Isso pode ser observado ao se analisar os valores de efeito de substituição alélica dos marcadores significativos (Tabela 5). Por exemplo, para CRa em BN, a seleção de genótipos com o marcador identificado no cromossomo 7 permite um incremento de 10,3% de comprimento de raiz (cm) em relação à média desse caráter, o que seria favorável para uma condição de baixa disponibilidade de nitrogênio. A informação desses efeitos, em conjunto com a informação fenotípica, pode ser usada para aumentar o ganho genético e a acurácia de seleção (SMITH; SIMPSON, 1986).

Portanto, a análise de associação pode ser muito efetiva em identificar a base e a arquitetura genética de caracteres quantitativos. Seguindo para o exercício de validação, os marcadores podem fornecer ferramentas de seleção indireta viáveis em um programa de melhoramento genético.

6. CONCLUSÃO

- As análises de GWAS evidenciaram associações significativas para os caracteres relacionados com eficiência no uso de nitrogênio em milho tropical, sendo que alguns SNP detectados localizam-se em regiões já reportadas na literatura, evidenciando a acurácia da análise e a importância dessas regiões genômicas.
- Dentre os principais processos biológicos identificados por meio da anotação funcional, está o controle e regulação da transcrição, detectado para todos os caracteres avaliados, e a síntese de GMP, enzima diretamente envolvida na disponibilização e reciclagem de N.
- Utilizando-se apenas os SNP identificados, não seria possível realizar seleção com acurácia para EUN. No entanto, esses marcadores podem auxiliar no processo seletivo ao serem usados como fatores de efeito fixo no modelo genético-estatístico.

REFERÊNCIAS

- ABDEL-GHANI, A. H. et al. Genotypic variation and relationships between seedling and adult plant traits in maize (*Zea mays* L.) inbred lines grown under contrasting nitrogen levels. **Euphytica**, v. 189, p. 123–133, 2013.
- ALVAREZ-BUYLLA, E. et al. MADS-box Gene Evolution Beyond Flowers: Expression in Pollen, Endosperm, Guard Cells, Roots and Trichomes. **The Plant Journal**, v. 24, p. 457-466, 2000.
- ATLIN, G. N.; FREY, K. J. Breeding crop varieties for low-input agriculture. **American Journal of Alternative Agriculture**, v. 2, p. 53-66, 2000.
- Bernardo R.; YU, J. Prospects for genome wide selection for quantitative traits in maize. **Crop Science**, v. 47, p. 1082-1090, 2007.
- BOOMSMA, C. R. Maize per plant and canopy-level morpho-physiological responses to the simultaneous stresses of intense crowding and low nitrogen availability. **Agronomy Journal**, v. 101, p. 1426–1452, 2009.
- BOUCHET, S. et al. Adaptation of Maize to Temperate Climates: Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with a Major Contribution of the Vgt2 (ZCN8) Locus. *Plos One*, v. 8, e71377, 2013.
- BRADBURY, P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. **Bioinformatics**, v. 23, p. 2633-2635, 2007.
- CECARELLI, S. Adaptation to low/high input cultivation. **Euphytica**, v.92, p.203-214, 1996.
- CHAO, D. et al. Sphingolipids in the Root Play an Important Role in Regulating the Leaf Ionome in *Arabidopsis thaliana*. **The Plant Cell**, v. 23, p. 1061–1081, 2011.
- CHUN, L. et al. Genetic analysis of maize root characteristics in response to low nitrogen stress. *Plant and Soil*, v. 276, p. 369-382, 2005.
- CIAMPITTI, I. A.; VYN, T. J. A comprehensive study of plant density consequences on nitrogen uptake dynamics of maize plants from vegetative to reproductive stages. **Field Crops Research**, v.121, p.02-18, 2011.
- CONAB – COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento de safra brasileira: grãos**, segundo levantamento, novembro 2016. Brasília: CONAB, 2016. 156 p.
- COOK, J. P. et al. Genetic Architecture of Maize Kernel Composition in the Nested Association Mapping and Inbred Association Panels. **Plant Physiology**, v. 158, n. 2, p. 824-834, 2012.

- CORSOUL, S. et al. Sphingolipid signalling in *Arabidopsis* guard cells involves heterotrimeric G proteins. **Nature**, v. 423, p. 651-654, 2003.
- COSTA, C. et al. Root morphology of contrasting maize genotypes. **Agronomy Journal**, v. 94, p. 96–101, 2002.
- CREN, M.; HIREL, B. Glutamine synthetase in higher plants: regulation of gene and protein expression from the organ to the cell. *Plant Cell Physiology*, v. 40, p. 1187-1193, 1999.
- DAMATTA, F. M. et al. Photosynthesis in coffee (*Coffea arabica* and *C. canephora*) as affected by winter and summer conditions. **Plant Science**, v. 128, n. 01, p. 43-50, 1997.
- DHANAPAL, A. P. et al. Genome-Wide Association Analysis of Diverse Soybean Genotypes Reveals Novel Markers for Nitrogen Traits. **The Plant Genome**, v. 8, n. 3, 2015.
- DOVALE, J. C. et al. Efeitos gênicos de caracteres associados à eficiência no uso de nitrogênio em milho. **Pesquisa Agropecuária Brasileira**, v. 47, n. 3, p. 385-392, 2012.
- DOVALE, J. C. et al. Melhoramento para eficiência no uso de nitrogênio. In: FRITSCHENETO, R.; BORÉM, A. **Melhoramento de plantas para condições de estresses abióticos**. Visconde do Rio Branco: Suprema, 2011. p. 81-99.
- DUARTE, A. P. Milho safrinha: características e sistema de produção. In: GALVÃO, J. C. C.; MIRANDA, G. V. **Tecnologias de produção do milho**. Viçosa: UFV, 2008. p. 109-138.
- DUARTE, A. P.; KAPPES, C. Evolução dos sistemas de cultivo de milho no Brasil. **Informações Agrônomicas**, n. 152, p. 15-18, 2015. INPI.
- DUBOIS, F. Glutamate dehydrogenase in plants: is there a new story for an old enzyme? **Plant Physiology and Biochemistry**, v. 41, p. 565-576, 2003.
- DUNN, T. M. et al. A post-genomic approach to understanding sphingolipid metabolism in *Arabidopsis thaliana*. **Annals of Botany**, v. 93, p. 483-497, 2004.
- EGHBALL, B.; MARANVILLE, J. W. Root development and nitrogen influx of corn genotypes grown under combined drought and N stress. **Agronomy Journal**, v. 85, p. 147–152, 1993.
- EISENBERG, D. Structure–function relationships of glutamine synthetases. **Biochimica et Biophysica Acta**, v. 1477, p. 122-145, 2000.
- EMBRAPA. **Cultivo do milho**. Disponível em: < <https://www.spo.cnptia.embrapa.br>>. Acesso em: 29 julho de 2016.
- EPSTEIN, E.; BLOOM, A. J. **Nutrição mineral de plantas: princípios e perspectivas**. Londrina: Editora Planta, 2006. 403 p.

- FARFAN, I. D. B. et al. Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. **Plos One**, 10(2):e0117737, 2015.
- FIDELIS, R. R. et al. Fontes de germoplasma de milho para estresse de baixo nitrogênio. **Pesquisa Agropecuária Tropical**, v. 3, p. 147-153, 2007.
- FRITSCHÉ-NETO, R.; BORÉM, A. **Melhoramento de plantas para condições de estresses abióticos**. Visconde do Rio Branco: Suprema, 2011. 250 p.
- FRITSCHÉ-NETO, R. et al. Herança de caracteres associados à eficiência de utilização do fósforo em milho. **Pesquisa Agropecuária Brasileira**, v. 45, p. 465-471, 2010.
- FRITSCHÉ-NETO, R. et al. Seleção genômica ampla e novos métodos de melhoramento do milho. **Ceres**, v. 59, p. 794-802, 2012.
- FRITSCHÉ-NETO, R.; BORÉM, A. **Plant Breeding for Abiotic Stress Tolerance**. 1th ed. Heidelberg: Springer, 2012. 172 p.
- GALLAIS, A.; HIREL, B. An approach to the genetics of nitrogen use efficiency in maize. **Journal of Experimental Botany**, v. 55, p. 295-306, 2004.
- GALVÃO, J. C. C.; TROGELLO, E.; PEREIRA, L. P. L. Milho segunda safra. In: BORÉM, A.; GALVÃO, J. C. C.; PIMENTEL, M. A. **Milho do plantio à colheita**. Viçosa: Ed. UFV, 2015. p. 207-223.
- GAUDIN, A. C. M. et al. Novel temporal, fine-scale and growth variation phenotypes in roots of adult-stage maize (*Zea mays* L.) in response to low nitrogen stress. **Plant, Cell and Environment**, v. 34, p. 2122-2137, 2011.
- GIÁCOMO, R. G.; PEREIRA, M.G.; GUARESCHI, R.F.; MACHADO, D.L. Atributos químicos e físicos do solo, estoques de carbono e nitrogênio e frações húmicas em diferentes formações vegetais. **Ciência Florestal**, v. 25, n. 3, p. 617-631, 2015.
- GILMOUR, A. R. et al. **ASReml user guide release 3.0**. Hemel Hempstead: VSN International, 2009. 398 p.
- GORE, M. A. et al. A First-Generation Haplotype Map of Maize. **Science**, v. 326, n. 5956, p. 1115-1117, 2009.
- GOWDA, M. et al. Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. **Theoretical and Applied Genetics**, v. 128, n. 10, p. 1957–1968, 2015.
- HORST, W. J. et al. Genotypic differences in nitrogen use-efficiency in crop plants. In: LYNCH, J. M., SCHEPERS, J. S., ÜNVER, I (eds). **Innovative Soil–Plant Systems for Sustainable Agricultural Production**. Paris: OECD Publishing, 2003. p. 75–92.

- IFA – INTERNATIONAL FERTILIZER INDUSTRY ASSOCIATION. **Assessment of fertilizer use by crop at the global level**. Paris: IFA, 2013. 10 p.
- IMEA – Instituto Mato-Grossense de Economia Agropecuária. **Custo de produção de soja e de milho-safras**: Mato Grosso. Cuiabá, 2016. Disponível em: <<http://www.imea.com.br>>. Acesso em: 30 de setembro de 2016.
- KANT, S. et al. Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. **Journal of Experimental Botany**, vol. 62, n. 4, p. 1499–1509, 2011.
- KNUEPPEL, S. et al. **Package ‘HapEstXXR’- Multi-Locus Stepwise Regression (MSR)**. 2015. 23 p.
- KORTE, A.; FARLOW, A. The advantages and limitations of trait analysis with GWAS: a review. **Plant Methods**, v. 9, p. 29, 2013.
- LAITY, J. H.; LEE, B. M.; WRIGHT, P. E. Zinc finger proteins: new insights into structural and functional diversity. **Current Opinion in Structural Biology**, v. 11, p. 39-46, 2001.
- LEA, P. J.; IRELAND, R. J. Plant amino acids. In: SINGH, B. K, ed. **Nitrogen metabolism in higher plants**. New York: Marcel Dekker Inc, 1999. p. 1-47.
- LI, H. et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. **Nature Genetics**, v. 45, n. 1, p. 43–50, 2013.
- LINKOHR, B. I. et al. Nitrate and phosphate availability and distribution have different effects on root system architecture of *Arabidopsis*. **Plant Journal**, v. 29, p.751–760, 2002.
- LIU, N. et al. Genome-Wide Association Study Identifies Candidate Genes for Starch Content Regulation in Maize Kernels. **Frontiers in Plant Science**, v. 7, article 1046, 2016.
- LIU, X. et al. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. **Plos One**, 12(2):e1005767, 2016.
- LYNCH, D. V.; DUNN, T. M. An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. **New Phytologist**, v. 161, p. 677-702, 2004.
- MACKAY, A. D.; BARBER, S. A. Effect of nitrogen on root growth of two corn genotypes in the field. **Agronomy Journal**, v. 78, p. 699–703, 1986.
- MAJEROWICZ, N. et al. Estudo da eficiência de uso do N em variedades locais e melhoradas de milho. **Revista Brasileira de Botânica**, v.25, p.129-136, 2002.
- MAMMADOV, J. et al. Combining powers of linkage and association mapping for precise dissection of QTL controlling resistance to gray leaf spot disease in maize (*Zea mays* L.). **BMC Genomics**, v. 16, n. 1, p. 1-16, 2015.

- MEDICI, L. O. et al. The influence of nitrogen supply on antioxidant enzymes in plant roots. **Functional Plant Biology**, v. 31, p. 1-9, 2004.
- MICHAELSON, L. V. et al. Plant sphingolipids: Their importance in cellular organization and adaption. **Biochimica et Biophysica Acta**, v. 1861, p. 1329–1335, 2016.
- MITI, F.; TONGOONA, P.; DERERA, J. S1 selection of local maize landraces for low soil nitrogen tolerance in Zambia. **African Journal of Plant Science**, v. 4, p.67-81, 2010.
- MONNEVEUX, P.; ZAIDI, P. H.; SANCHEZ, C. Population Density and Low Nitrogen Affects Yield-Associated Traits in Tropical Maize. **Crop Science**, vol. 45, p. 535-545, 2005.
- MORAIS, P. P. P. Eficiência no uso de nitrogênio ao longo dos estádios fenológicos e utilização de dados públicos na seleção genômica de milho tropical. 2016. 78 p. Tese (Doutorado em Genética e Melhoramento. Universidade Federal de Viçosa, Viçosa, 2016.
- MURRAY, M.; THOMPSON, W. F. Rapid isolation of high molecular weight plant DNA. **Nucleic Acids Research**, v. 8, n. 19, p. 4321-4326, 1980.
- PACE, J. et al. Genome-wide association analysis of seedling root development in maize (*Zea mays* L.). **BMC Genomics**, v. 16, n. 1, p. 16-47, 2015.
- PECCHIO, M. S. Sistemas de produção do milho safrinha no Médio Vale Paranapanema, estado de São Paulo. In: SEMINÁRIO NACIONAL DE MILHO SAFRINHA, 13., 2015. **Anais...** Maringá: UEM/IAPAR/EMAER, 2015. p. 592-602.
- PEDROSA, A. W. **Eficiência da adubação nitrogenada ao consórcio entre cafeeiro e *Brachiaria brizantha***. 2013. 74 p. Tese (Doutorado em Fitotecnia) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2013.
- PEIFFER, J. A. et al. The genetic architecture of maize height. **Genetics**, v. 196, n. 4, p. 1337–1356, 2014.
- PENG, Y.; YU, P.; LI, X.; LI, C. Determination of the critical soil mineral nitrogen concentration for maximizing maize grain yield. **Plant and Soil**, v. 372, p. 41-51, 2013.
- PETROBRAS. Petróleo Brasileiro S.A. **Análise do Mercado de Fertilizantes**, 2009. Disponível em: <<http://www.cetem.gov.br/agrominerais/livros/AnaliseMercado.pdf>>. Acesso em 03 de setembro de 2016.
- PNUCLI, L. et al. Isolation of the tomato AGAMOUS Gene *TAG1* and analysis of its homeotic role in transgenic plants. **Plant Cell**, v. 6, p. 163-173, 1994.
- R CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2014.

- RAFALSKY, J. A. Association genetics in crop improvement. **Current Opinion in Plant Biology**, v. 13, p. 174-180, 2010.
- REICHMANN, J., MEYEROWITZ, E. MADS-domain Proteins in Plant Development. **Biology Chemistry**, v. 378, p. 1079-1101, 1997.
- RESENDE, M. D. Seleção Genômica Ampla (GWS). In: BORÉM, A.; FRITSCHÉ-NETO, R. ed. **Biotecnologia aplicada ao melhoramento de plantas**. Visconde de Rio Branco: Suprema, 2013. 336 p.
- RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v. 37, n. 3, p. 182–194, 2007.
- ROMAY, M. et al. Comprehensive genotyping of the USA national maize inbred seed bank. **Genome Biology**, v.14, R55, 2013.
- SANT'ANNA, G. C. **Relação da estrutura genética com a eficiência no uso de nitrogênio e com a acurácia de GWAS em milho tropical**. 2015. 77 p. Tese (Doutorado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa, 2015.
- SATTELMACHER, B.; KLOTZ, F.; MARSCHNER, H. Influence of the nitrogen level on root growth and morphology of two potato varieties differing in nitrogen acquisition. **Plant Soil**, v. 123, p. 131–137, 1990.
- SAUCEDO-GARCIA, M. et al. MPK6, sphinganine and the LCB2a gene from serine palmitoyltransferase are required in the signaling pathway that mediates cell death induced by long chain bases in *Arabidopsis*. **New Phytologist**, v. 191, p. 943-957, 2011.
- SCHNABLE, P. S. et al. The B73 Maize Genome: Complexity, Diversity, and Dynamics . **Science**, v. 326, p. 1112-1115, 2009.
- SINCLAIR, T. R.; GARDNER, F. P. Environment limits to plant production. In: SINCLAIR, T. R.; GARDNER, F. P. **Principles of ecology in plant production**. New York: CAB International, 1998. p. 63-78.
- SMITH, C. SIMPSON, S. P. The use of genetic polymorphisms in livestock improvement. **Journal of Animal Breeding and Genetics**, v. 103, p. 205-217, 1986.
- SOUZA, L. V. et al. Combining ability of maize grain yield under different levels of environmental stress. **Pesquisa Agropecuária Brasileira**, v. 44, p. 1297-1303, 2009.
- SOUZA, L. V. et al. Genetic control of grain yield and nitrogen use efficiency in tropical maize. **Pesquisa Agropecuária Brasileira**, v.43, p.1517-1523, 2008.
- STELPFLUG, S. C. et al. An expanded maize gene expression atlas based on RNA-sequencing and its use to explore root development. **The Plant Genome**, v. 9, n. 1, p. 1-16, 2016.

- STEPONKUS, P. L.; LYNCH, D. Freeze/thaw-induced destabilization of the plasma membrane and the effects of cold acclimation. **Journal of Bioenergetics and Biomembranes**, v. 21, p. 21-41, 1989.
- TASCA, F. A. et al. Volatilização de amônia no solo após a aplicação de ureia convencional ou com inibidor de urease. **Revista Brasileira de Ciência do Solo**, v. 35, p.493-502, 2011.
- THEISSEN, G.; KIM, J.; SAEDLER, H. Classification and Phylogeny of the MADS-box Multigene Family Suggest Defined Roles of MADS-box Gene Subfamilies in the Morphological Evolution of Eukaryotes. **Journal Molecular Evolution**, v. 43, p. 484-516, 1996.
- TIAN, F. et al. Genome-wide association study of leaf architecture in the maize nested association mapping population. **Nature Genetics**, v. 43, p. 159-162, 2011.
- TOALDO, D. et al. Selection in early generations and the occurrence of heterosis for the character root distribution. **Euphytica**, v. 190, p. 335–344, 2013.
- UEMURA, M.; JOSEPH, R. A.; STEPONKUS, P. L. Cold acclimation of *Arabidopsis thaliana* (effect on plasma membrane lipid composition and freeze-induced lesions). **Plant Physiology**, v. 109, p. 15-30, 1995.
- UNTERSEER, S. et al. A powerful tool for genome analysis in maize: development and evaluation of the high density 600k SNP genotyping array. **BMC genomics**, v. 15, n. 1, p. 823, 2014.
- VANRADEN, P. M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, v. 91, n. 11, p. 4414–4423, 2008.
- VARSHNEY, R. K. et al. Genomics and physiological approaches for root trait breeding to improve drought tolerance in Chickpea (*Cicer arietinum* L.) In: OLVEIRA, A. C.; VARSHNEY, R. K. Root Genomics. Heidelberg: Springer, 2011. cap. 10, p. 233-250.
- WANG, Q. et al. A SUPER Powerful Method for Genome Wide Association Study. **Plos One**, v. 9, n. 9, e107684, 2014.
- WANG, Y. Response of root morphology to nitrate supply and its contribution to nitrogen accumulation in maize. **Journal of Plant Nutrition**, v. 27, p. 2189– 2202, 2004.
- WHITT, S. R.; BUCKLER, E. S. Using natural allelic diversity to evaluate gene function. **Methods in Molecular Biology**, v. 236, p. 123-140, 2003.
- WORRALL, D. et al. Involvement of sphingosine kinase in plant cell signaling. **The Plant Journal**, v. 56, p. 64-72, 2008.

- WU, Y. et al. Low-nitrogen stress tolerance and nitrogen agronomic efficiency among maize inbreds: comparison of multiple indices and evaluation of genetic variation. **Euphytica**, Amsterdam, v. 180, p. 281-290, 2011.
- XIE, L. J. et al. Unsaturation of very-long-chain ceramides protects plant from hypoxia-induced damages by modulating ethylene signaling in Arabidopsis, **Plos Genetics**, v. 11, e1005143, 2015.
- XU, Y. F. et al. Mapping QTLs for yield and nitrogen-related traits in wheat: influence of nitrogen and phosphorus fertilization on QTL expression. **Theoretical and Applied Genetics**, v. 127, n. 1, p. 59-72, 2014.
- XUE, Y. et al. Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. **Theoretical and Applied Genetics**, v. 126, p. 2587–2596, 2013.
- ZHANG, X. et al. Identification of Major QTL for Waterlogging Tolerance Using Genome-Wide Association and Linkage Mapping of Maize Seedlings. **Plant Molecular Biology Reporter**, v. 31, p. 594–606, 2013.
- ZONTA, E. et al. O sistema radicular e suas interações com o ambiente edáfico. In: FERNANDES, M.S., ed. **Nutrição mineral de plantas**. Viçosa, MG: Sociedade Brasileira de Ciência do Solo, 2006. p. 7-52.

ANEXOS

ANEXO A. Análise de solo da área experimental do Departamento de Genética da ESALQ/USP em Piracicaba – SP

Determinações			Unidades	Amostras ¹				
-	-	-	-	1	2	3	4	5
Ph	CaCl ₂	CaCl ₂ 0,01mol/L	-	4,7	5,3	5,1	4,9	5,3
P	Resina trocadora de íons	mg/dm ³		39	46	57	53	47
K	Resina trocadora de íons	mmolc/dm ³		1,8	3,8	5	5,0	3,5
Ca	Resina trocadora de íons	mmolc/dm ³		24	43	27	18	28
Mg	Resina trocadora de íons	mmolc/dm ³		8	22	9	6	9
Cl	Resina trocadora de íons	mmolc/dm ³		2	2	2	2	2
H+Al	Alumínio trocável KCL 1mol/L	mmolc/dm ³		31	22	28	38	25
SB	Acidez Potencial – SMP	mmolc/dm ³		33,8	68,8	41	29,5	40,5
CTC	-	mmolc/dm ³		64,8	90,8	69	67,5	65,5
V	-	%		52	76	59	44	62
m	-	%		6	3	5	6	5
M.O.	-	g/dm ³		18	19	18	21	21

¹ Data de emissão: 18/09/2014

ANEXO B. Dados meteorológicos do Ambiente 3 (Piracicaba 2015)

MÊS	PRECIPIT AÇÃO	UMIDADE RELATIVA	VENTO MÁXIMO	VENTO MÍNIMO	T ^a MÁXIMA	T ^a MÍNIMA	T ^a MÉDIA	EVAPO RAÇÃO
	mm	%	m/s	km/h	°C	°C	°C	mm
JAN	104.1	72	10.2	6.1	33.7	21.6	27.7	7.28
FEV	248.2	84	10.5	4.9	32.2	20.9	26.5	5.68
MAR	134.3	86	7.6	4.6	30.1	19.9	25.0	5.08
ABR	12.8	79	8.7	5.9	29.7	17.5	23.6	5.98
MAI	86.7	83	7.3	5.0	26.3	14.9	20.6	4.71
JUN	3.8	78	6.8	5.8	26.5	13.7	20.1	5.45