



VIVIANA RAMÍREZ RÍOS

**ASSOCIAÇÃO DE MARCADORES-QTL COM
CARACTERÍSTICAS DE TOLERÂNCIA AO DÉFICIT HÍDRICO EM
PLÂNTULAS DE MILHO**

**LAVRAS - MG
2024**

VIVIANA RAMÍREZ RÍOS

**ASSOCIAÇÃO DE MARCADORES-QTL COM CARACTERÍSTICAS DE
TOLERÂNCIA AO DÉFICIT HÍDRICO EM PLÂNTULAS DE MILHO**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia vegetal, área de concentração em Biotecnologia vegetal, para a obtenção do título de Doutora.

Profa. Dra. Édila Vilela de Resende Von Pinho
Orientadora
Profa. Dra. Heloisa Oliveira Dos Santos
Coorientadora

**LAVRAS - MG
2024**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Ramírez Ríos, Viviana.

Associação de marcadores-QTL com características de
tolerância ao déficit hídrico em plântulas de milho / Viviana
Ramírez Ríos. - 2024.

62 p. : il.

Orientador(a): Édila Vilela de Resende Von Pinho.

Coorientador(a): Heloísa Oliveira dos Santos.

Tese (doutorado) - Universidade Federal de Lavras, 2024.

Bibliografia.

1. Mapeamento genético. 2. Microssatélites. 3. Déficit hídrico.
I. Von Pinho, Édila Vilela de Resende. II. dos Santos, Heloísa
Oliveira. III. Título.

VIVIANA RAMÍREZ RÍOS

**ASSOCIAÇÃO DE MARCADORES-QTL COM CARACTERÍSTICAS DE
TOLERÂNCIA AO DÉFICIT HÍDRICO EM PLÂNTULAS DE MILHO**

**QTL-MARKER ASSOCIATIONS WITH CHARACTERISTICS OF TOLERANCE
TO WATER DEFICIT IN CORN SEEDLING**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia vegetal, área de concentração em Biotecnologia vegetal, para a obtenção do título de Doutora.

APROVADA em 28 de fevereiro de 2024.

Dra. Cláudia Teixeira Guimarães	EMBRAPA
Dra. Renata Silva Mann	UFS
Dra. Heloísa Oliveira dos Santos	UFLA
Dr. José Maria Villela Pádua	UFLA

Profa. Dra. Édila Vilela de Resende Von Pinho
Orientadora

**LAVRAS - MG
2024**

A Deus por não soltar a minha mão na tentativa.

À minha família por ser minha força motriz na vida e na luta.

A todos aqueles que me ajudaram e me acompanharam

no processo de formação acadêmica.

A minha pessoa por ter persistido na realização do meu sonho

Com todo o meu amor

DEDICO

AGRADECIMENTOS

Primeiramente agradeço a Deus, que sempre me deu fortaleza e discernimento.

Aos meus pais, Glória Elsy e Gustavo Alberto por me apoiar sempre em cada passo que dou na vida. Sem o amor de vocês, confiança, presença e suporte mesmo a distância, essa conquista não seria possível. Obrigada por me ensinarem a mim como sonhar, fortalecer, persistir e confiar nas minhas capacidades.

Às minhas irmãs Leidy e Cindy, por serem as minhas amigas e por me apresentar aos meus grandes amores, os meus sobrinhos Laura, Isabella, Alejandro e Celeste. Eles são a minha alegria diária e por eles buscarei sempre ser uma pessoa melhor.

À minha avó Luz Estella (*in memoriam*), por me ensinar sobre a nobreza, a fé e por sempre ter confiado em mim.

Aos meus tios Arnobis e Luz Dary por ter me apoiado e ajudado para continuar com os meus estudos quando para os meus pais não era possível me ajudar, sem vocês eu não teria chegado tão longe.

À Universidade Federal de Lavras e ao Programa de Pós-graduação em Biotecnologia Vegetal, pela oportunidade e pelo incentivo.

Ao programa de Parcerias para Educação e Capacitação (PAEC), conformado pela Organização dos Estados Americanos (OEA) e o Grupo de Cooperação Internacional das Universidades Brasileiras (GCUB), por me permitir a candidatura pela bolsa de estudos de doutorado.

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

À minha orientadora, profa. Dra. Édila Vilela, por permitir e me auxiliar na realização dessa pesquisa. Pelos ensinamentos, por confiar em meu trabalho e sobretudo por ser essa mulher incrível que sempre me deu suporte emocional e acadêmico da forma mais humana e profissional.

Aos professores da Pós-graduação, pela contribuição na minha formação técnica e cidadã. Em especial à professora Heloisa Oliveira dos Santos, pela constante presença, coorientação e pelos ensinamentos.

Aos membros da banca examinadora Pesquisadora Dra. Claudia Teixeira Guimarães, Profa. Dra. Renata Silva Mann, Prof. Dra. Heloisa Oliveira dos Santos, e Prof. Dr. José Maria Villela Pádua, pelas valiosas contribuições.

Aos secretários e funcionários do Laboratório Central de Biologia Molecular, Fabrício, Heliete e Ricardo; e aos funcionários do Departamento de Sementes, Jaqueline e Geraldo, por todo o suporte e atenção.

Às minhas amigas de colégio Yuly Ospina e Catalina Echeverry, pela amizade incondicional e sólida construída durante esses 20 anos de amizade, por serem o meu suporte emocional e o meu porto seguro sempre.

Ao meu professor e colega Javier Correa Alvarez, por sempre estar disposto a me ensinar, me guiar e me dar suporte acadêmico durante toda a minha vida profissional.

Aos amigos que fiz em Lavras, Janet, Samanda, Dayana, Karen, Nataly, Andrea, Mayra, Zulma, Robert, Miguel, Hector, Denisse e especialmente ao Alector Emanuel, por me acompanharem durante toda essa caminhada e serem meu suporte e carinho em todos os momentos, mais que amigos vocês foram a minha família aqui no Brasil e fizeram dessa caminhada mais leve e divertida.

Aos estagiários e bolsistas de iniciação científica, por todo auxílio e dedicação na condução dos experimentos.

A todos que de alguma forma fizeram parte dessa conquista, sem vocês essa conquista não seria possível.

Muito obrigada!

RESUMO

O milho é uma das culturas mais importantes no Brasil, mas o seu potencial de rendimento é severamente influenciado pelo déficit hídrico, devido às constantes mudanças climáticas, principalmente na segunda safra, safrinha, mas também não incomum na primeira safra, com prejuízos aos agricultores. A identificação de cultivares de milho tolerantes ao déficit hídrico em estádios iniciais de plântulas, pode ser uma opção em processos de seleção, em programas de melhoramento. Para esta finalidade é necessário identificar regiões do genoma que possam ser usadas como marcadores de características de tolerância ao déficit hídrico. Assim, experimentos foram conduzidos para identificar regiões associadas à tolerância ao déficit hídrico em genótipos de milho classificados no estádio de plântulas quanto a esta característica. Para mapear regiões relacionadas à tolerância ao déficit hídrico (locos de características quantitativas, QTL), foram utilizados 205 indivíduos pertencentes a duas populações segregantes F₂ obtidas a partir do cruzamento de duas linhagens parentais P₁ (tolerante) e P₂ (suscetível). Sementes desses indivíduos foram submetidas ao déficit hídrico induzido por PEG 6000 (-0.6MPa), e a germinação avaliada 5 dias após semeadura e na sequência classificadas por nível de tolerância ao déficit hídrico em tolerantes, intermediárias e suscetíveis; em função do desenvolvimento das plântulas. Para a genotipagem, marcadores microssatélites (SSR) localizados próximos aos QTLs para esta característica, foram selecionados de acordo com os dados da literatura. O DNA genômico foi extraído pelo método CTAB e os produtos de PCR foram separados por eletroforese capilar. As populações analisadas segregaram como 1 (tolerante): 2 (intermediário):1 (suscetível), indicando um modelo de segregação mendeliana. As análises dos dados de cada marcador SSR nas progênies foram feitas nos softwares *Cervus* e *InfoGen*, enquanto a análise de associação foi feita por meio do programa *InfoGen*. Associações significativas dos dados fenotípicos e genotípicos, foram detectadas para os marcadores p-bnlg1208 e phi027. Utilizando-se ferramentas de bioinformática e análises *in silico* nos dois locos com associação à tolerância ao déficit hídrico, foram identificados seis genes que podem ser alvo de futuros estudos que possuem funções relacionadas a esta característica, sendo potenciais marcadores para seleção precoce assistida por marcadores (MAS) em programas de melhoramento de milho.

Palavras chave: mapeamento genético, microssatélites, déficit hídrico, tolerância, milho.

ABSTRACT

Maize is one of the most important crops in Brazil, but its yield potential is severely influenced by the water deficit, due to constant climate changes, mainly in the second harvest, but also not uncommon in the first harvest, with losses to farmers. Selective breeding can offer an opportunity for farmers through identifying maize types that are tolerant to drought in the initial seedling stage. For this purpose, it was necessary to identify regions of the genome that can be used as markers for drought tolerance. Thus, experiments were conducted to identify regions associated with drought tolerance in maize genotypes classified at the seedling stage for this characteristic. To map regions related to tolerance to water deficit (quantitative trait locos, QTL), 205 individuals belonging to two F₂ segregating populations obtained from the crossing of two parental lines P1 (tolerant) and P2 (intolerant) were used. Seeds from these individuals were subjected to water deficit induced by PEG 6000 (-0.6MPa), and germination was evaluated 5 days after sowing and subsequently classified by level of tolerance to water deficit (tolerant, intermediate and intolerant) depending on development of the seedlings. For genotyping, SSR microsatellite markers located close to QTLs for this trait were selected according to literature data. Genomic DNA was extracted by the CTAB method and PCR products were separated by capillary electrophoresis. The analyzed populations segregated as 1 (tolerant): 2 (intermediate): 1 (intolerant), indicating a Mendelian segregation model. Data analyzes of each SSR marker in the progenies were carried out using the *Cervus* and *InfoGen* software, while the association analysis was carried out using the *InfoGen* program. Significant associations of phenotypic and genotypic data were detected for markers bnlg1208 and phi027. Using bioinformatics tools and *in silico* analyzes at the two locos associated with tolerance to drought six genes were identified that could be the target of future studies, being potential markers for early selection assisted markers (MAS) in corn breeding programs and that have functions related to drought tolerance.

Keywords: genetic mapping, microsatellites, water stress, tolerance, maize

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

Tabela 1- Marcadores SSR polimórficos nas análises em bulks por eletroforese em gel de poliacrilamida.	29
Tabela 2- Tamanho e numero de alelos polimórficos determinados pelo sistema de eletroforese capilar.	32
Tabela 3- Medidas de diversidade genética por loco.	35
Tabela 4- frequências genotípicas relativas por loco.....	36
Tabela 5- Análises da variância entre os genótipos por loco.	37
Tabela 6- Dados estatísticos da comparação entre médias resultado da ANOVA.....	38
Tabela 7- Frequências absolutas por nível de tolerância ao déficit hídrico.....	41
Tabela 8- Genes candidatos identificados nas proximidades 100KB dos locos bnlg1208 e phi027.....	46

LISTA DE FIGURAS

Figura 1- Segregação da característica fenotípica na população F _{2:3}	20
Figura 2- Esquema ilustrativo da metodologia de fenotipagem e obtenção da população segregante.	21
Figura 3- Diagrama ilustrativo da metodologia usada para a extração do DNA e as análises em bulks.....	23
Figura 4- Framework computacional simplificado baseado nas análises bioinformáticas que foram empregadas para a identificação de genes associados aos locos relacionados com a tolerância ao déficit hídrico em Milho.	28
Figura 5-Resultados de polimorfismo dos marcadores SSR nas análises em bulks em géis de poliacrilamida.	30
Figura 6- Resultados das análises de polimorfismo entre os parentais P1 (tolerante) e P2 (suscetível), no sistema de eletroforese capilar.	32
Figura 7- Eletroferograma e análise de dados gerados pelo sistema de eletroforese capilar TapeStation 4150 (Agilent Technologies) para o caso de dois alelos muito próximos.	33

Figura 8- Comparação de médias entre os genótipos por loco (ANOVA).....	39
Figura 9- Análise de regressão linear. Efeito médio da substituição alélica	40
Figura 10- Representação gráfica dos cromossomos 5 e 9 mostrando os genes encontrados nas proximidades dos locos p-blng1208 e phi027.	47

SUMÁRIO

1. INTRODUÇÃO	13
2. MATERIAL E MÉTODOS	19
2.1 OBTENÇÃO DAS POPULAÇÕES SEGREGANTES (FENOTIPAGEM)	19
2.2 GENOTIPAGEM DA POPULAÇÃO F₂ SEGREGANTE	21
2.2.1. EXTRAÇÃO DE DNA E SELEÇÃO DOS MARCADORES SSR	21
2.2.2. GENOTIPAGEM DA POPULAÇÃO F ₂ POR ELETROFORESE CAPILAR	23
2.3. VARIABILIDADE GENÉTICA DOS MARCADORES SSR	24
2.4. ANÁLISES DE ASSOCIAÇÃO ENTRE OS SSR POLIMÓRFICOS E A TOLERÂNCIA AO DÉFICIT HÍDRICO (ASSOCIAÇÃO MARCADOR-QTL)	26
2.5. IDENTIFICAÇÃO IN SILICO DE GENES ASSOCIADOS AOS LOCOS SSR bnlg1208 E phi027	27
3. RESULTADOS E DISCUSSÃO.....	28
4. CONCLUSÕES.....	51
REFERÊNCIAS.....	52
MATERIAL ANEXO.....	59

1. INTRODUÇÃO

O milho (*Zea mays L.*) é um dos cereais mais cultivados no mundo, com mais de 1501.8 milhões de toneladas produzidas em 2021/2022, com previsão de crescimento da produção em 6% para a safra 2023/2024 (USDA, 2023) também há previsão de aumento na produtividade, resultante da melhoria de tecnologias e das práticas de cultivo, em particular nos principais países produtores (FAO, 2023).

O Brasil, como o terceiro produtor e o maior exportador mundial de milho, vem desenvolvendo novas tecnologias e estratégias que possam viabilizar o estabelecimento da cultura para o aumento dos rendimentos (FORBES AGRO, 2023). Sabe-se que a evolução recente da produção e da produtividade de milho no Brasil é resultado das pesquisas desenvolvidas pela comunidade acadêmica e principais centros de pesquisa em agricultura no país, que tem permitido revolucionar a agricultura brasileira por meio de novas cultivares adaptadas às diferentes regiões e técnicas de cultivo, sob condições tropicais (KLEIN; LUNA, 2022). No entanto, a produção de milho ainda é limitada devido aos estresses bióticos e abióticos, sendo o déficit hídrico o principal fator que limita a produtividade de grãos, principalmente na segunda safra (safrinha), mas também não incomum na primeira safra, trazendo prejuízos aos agricultores (VON PINHO, 2021).

A tolerância ao déficit hídrico no milho é uma característica complexa influenciada por múltiplos genes e suas interações. As plantas têm a capacidade de detectar e responder a condições estressantes alterando sua expressão gênica a fim de induzir modificações fisiológicas e bioquímicas. Alguns dos genes e mecanismos principais associados à tolerância ao estresse hídrico incluem: Fatores de transcrição (CAO *et al.*, 2024; HE *et al.*, 2020), genes da via dependente de ABA (KIZIS; PAGÈS, 2002), proteínas LEA (Abundante Embriogênese Tardia) (ZHANG *et al.*, 2023), aquaporinas (HASAN *et al.*, 2017), genes *ZmSnRK2* (proteína quinase 2 não fermentadora de sacarose 1) (LONG *et al.*, 2021), genes *ZmCAT* (catalase) e *ZmSOD* (superóxido dismutase) (TALAAT; IBRAHIM; SHAWKY, 2022) e genes *ZmPIP* (proteína intrínseca da membrana plasmática) (AUOBI; RAHMANI, 2017).

Para reduzir a perda de água, as plantas empregam várias estratégias, como a remoção de espécies reativas de oxigênio (ROS), o acúmulo de agentes de osmorregulação e o fechamento dos estômatos. A regulação da abertura estomática nas plantas está intimamente

ligada à presença do ácido abscísico (ABA), que está sujeito à modulação dependendo da gravidade do estresse sofrido pela planta (HU; XIONG, 2014).

O acúmulo de ácido abscísico (ABA) pode dar origem a sinais de respostas adicionais quando o estresse hídrico está presente como a indução de diferentes sistemas de defesa antioxidantes, tais como enzimas superóxido dismutase, peroxidase e catalase e antioxidantes não enzimáticos como o ácido ascórbico, glutatona e carotenoides. O hormônio ABA ainda induz a atuação de vários componentes importantes na defesa antioxidante, como cálcio (Ca^{2+}), calmodulina, proteína quinase ativada por nitrogênio (*MAPK*) e proteína quinase dependente de cálcio/calmodulina (*CcaMK*) (CHRISTY SANTOS, 2016). Assim, o ABA desempenha um papel fundamental na regulação da condutância estomática e é recrutado para iniciar respostas adaptativas ao déficit hídrico (CAO *et al.*, 2024; HE *et al.*, 2020).

Nas plantas, a presença de fatores de transcrição (TFs) é vital para facilitar as respostas ao estresse por meio de redes de sinais. Esses fatores de transcrição têm a capacidade de se ligar a elementos cis-reguladores que estão localizados nos promotores, regulando assim a taxa de transcrição dos genes. Dentro do sistema de resposta ao estresse das plantas, os fatores de transcrição desempenham um papel fundamental nos mecanismos que governam a resposta ao estresse das plantas, incluindo *AP2/EREB* (Apetala2/ proteínas de ligação a elementos responsivos ao etileno), *bZIP* (domínio zíper de leucina básica) e *DREB2A* (elemento de ligação responsivo à desidratação), *ZmNAC*, todos associados ao estresse hídrico nas plantas (CAO *et al.*, 2024; HE *et al.*, 2020). Os outros grupos de genes como os codificantes para proteínas LEA, regulam o desenvolvimento do grão e os processos de desidratação, especialmente membros da subfamília LEA 4 (ZHANG *et al.*, 2023). As aquaporinas facilitam o transporte de água através das membranas celulares e parecem estar envolvidos na regulação da captação e distribuição de água em resposta ao estresse hídrico. Os genes *ZmSnRK2* fazem parte da via de sinalização ABA (LONG *et al.*, 2021).

Os genes *ZmCAT* (catalase) e *ZmSOD* (superóxido dismutase), codificam enzimas antioxidantes, como a catalase e a superóxido dismutase e, fazem parte do grande sistema antioxidante enzimático que são essenciais na eliminação de espécies reativas de oxigênio (ROS) produzidas sob condições de estresse em plantas e mantendo a integridade da membrana celular (TALAAT; IBRAHIM; SHAWKY, 2022).

Os genes *ZmPIP* (proteína intrínseca da membrana plasmática) que são semelhantes às aquaporinas. Os genes *PIP* codificam proteínas envolvidas no transporte de água através das membranas celulares e podem contribuir para a tolerância ao déficit hídrico (TALAAT; IBRAHIM; SHAWKY, 2022).

Embora tenham-se grupos de genes bem estudados em quanto a sua relação com a tolerância ao déficit hídrico, existem outros grupos de genes que ainda não tem uma função conhecida ou que tem escassos estudos publicados até o momento.

É importante notar que a interação de múltiplos genes, bem como de fatores ambientais, desempenham um papel significativo na determinação da tolerância geral ao déficit hídrico em milho, além de ter em conta que a pesquisa neste campo ainda está em andamento e novas descobertas estão em constante surgimento (LIU *et al.*, 2020a). No entanto, novas descobertas que permitam contribuir com a compreensão da base genética da tolerância ao estresse hídrico em milho, precisam do desenvolvimento de cultivares tolerantes em solos com condições desfavoráveis o que depende a sua vez dos avanços nas técnicas e metodológicas de análises genéticas e moleculares usadas para tal finalidade.

Um dos maiores desafios em programas de melhoramento de milho é a seleção de cultivares com tolerância ao déficit hídrico, em função da natureza quantitativa dessa característica. Embora em alguns programas de melhoramento de milho sejam utilizados marcadores moleculares para o enriquecimento do genoma e geração de novas cultivares com a característica de interesse, há necessidade urgente de empregar genotipagem econômica e rápida para a identificação de genótipos de milho tolerantes ao déficit hídrico (ROMERO *et al.*, 2017).

Em milho há substancial variabilidade genética, que pode ser explorada para o desenvolvimento de novas cultivares superiores. Assim, a análise da diversidade genotípica pode fornecer informações para a identificação de fontes de alelos importantes, que junto com dados fenotípicos, pode resultar numa possível introdução de genes de interesse por meio de técnicas biotecnológicas (PEER *et al.*, 2022). Neste contexto, métodos de fenotipagem e genotipagem e novas tecnologias podem ser desenvolvidas, com a finalidade de compreender as interações genótipo-fenótipo e, a estrutura genética das características quantitativas que ainda não são completamente compreendidas (ROMERO *et al.*, 2017).

Os métodos tradicionais de detecção de regiões associadas a características quantitativas requerem fenotipagem e genotipagem de grandes populações, tornando-os caros e de baixo poder. Técnicas como a análise em Bulks ou análise segregante agrupada (BSA) é particularmente especial para a genotipagem de marcadores, para características quantitativas

em populações grandes usando a exploração da tolerância ao déficit hídrico em plantas de milho (QUDRAT *et al.*, 2018). Os bulks são amostras constituídas pela combinação de DNA de várias plantas individuais de uma mesma população e com o mesmo fenótipo em uma única amostra. A estratégia da composição de bulks é originalmente baseada na mistura em iguais quantidades de DNA genômico extraído de plantas individuais, que geram resultados consistentes com a análise individual de cada DNA, permitindo filtrar a quantidade inicial de marcadores moleculares analisados (EMBRAPA CLIMA TEMPERADO, 2007).

Os microssatélites ou repetições de sequência simples (SSRs), são marcadores moleculares de muita utilidade na genotipagem associada a características como a tolerância ao déficit hídrico. Os SSR, são repetições de até 100 vezes de sequências simples de 1–8 pares de bases. Esses elementos estão presentes tanto em regiões codificantes como não codificantes de todos os genomas eucarióticos e procarióticos estudados até o momento, estando presentes até mesmo no cloroplasto e no DNA mitocondrial (GARRIDO-CARDENAS; MESA-VALLE; MANZANO-AGUGLIARO, 2018). Em milho, estima-se que ocorra um microssatélite a cada 58 Kb (SALGADO, 2005). Devido a esta estrutura repetitiva, os SSRs frequentemente sofrem mutações, principalmente devido a erros de DNA polimerase, que envolvem a adição ou subtração de uma unidade repetida. Assim, as sequências SSR são altamente polimórficas e podem ser facilmente utilizadas para detecção de variação alélica dentro de populações (MASON, 2015). Os SSRs têm sido um dos marcadores moleculares mais utilizados para a genotipagem de plantas devido aos seus altos níveis de polimorfismo, ampla distribuição nos genomas, codominância, herança mendeliana, alta variabilidade, natureza multialélica, altamente reprodutíveis e baratos (GARRIDO-CARDENAS; MESA-VALLE; MANZANO-AGUGLIARO, 2018; MASON, 2015; 2018; SHARMA *et al.*, 2018).

A genotipagem de SSR para a busca de associações com características quantitativas, envolve o desenho ou síntese de *primers* para amplificar sequências de SSR do DNA genômico extraído, seguido pela amplificação da região repetida de SSR usando reação em cadeia da polimerase e subsequente visualização dos produtos de DNA resultantes, geralmente usando eletroforese em gel ou eletroforese capilar (MASON, 2015; NAEEM, 2014). Algumas das limitantes para a genotipagem de populações de interesse são o desenho dos *primers* específicos que flanqueiam as regiões altamente repetitivas como os microssatélites SSR e a técnica de separação dos produtos das amplificações dessas regiões repetidas. No entanto, para o caso do milho, a primeira limitante foi vencida graças ao projeto MaizeGDB (<https://www.maizegdb.org/>) que possui mais de 2000 *primers* SSR distribuídos nos 10 cromossomos de milho e estão completamente disponíveis para o público (SALGADO, 2005).

Para o caso do sistema de visualização das regiões repetidas, embora existam novas tecnologias de visualização como a eletroforese capilar é uma área que continua em constante atualização para proporcionar resultados altamente confiáveis e reprodutíveis.

A eletroforese capilar é uma técnica empregada para separar fragmentos de DNA e outras moléculas, com base em seu tamanho e carga. Usa uma estrutura polimérica linear chamada capilar, normalmente é muito estreito e feito de sílica fundida, permitindo uma separação de alta resolução. O capilar é preenchido com uma solução tampão na qual é injetada a amostra de DNA previamente marcado com um intercalante de DNA fluorescente. Posteriormente, um campo elétrico é aplicado através do capilar, fazendo com que os fragmentos de DNA migrem por tamanho e carga, fazendo possível a sua separação. O sistema de detecção captura a fluorescência emitida pela molécula e gera um eletroferograma, que é uma representação gráfica da separação, permitindo caracterizar a molécula por tamanho (em pares de bases) e por qualidade usando como referência os picos *lower* e *upper* do eletroferograma (MITCHELSON; CHENG, 2008; RIGHETTI; GELFI; D'ACUNTO, 2002; TAVARES, 1997). Após a análise dos eletroferogramas, o processo da genotipagem de populações de interesse, requer de análises genéticas baseadas nas frequências alélicas e genóticas encontradas nas populações de estudo (MASON, 2015).

Uma série de testes de diversidade genética podem ser utilizados para a validação das diferenças e semelhanças genéticas entre as populações além do tipo de herança. Basicamente, eles são estimados baseados na proporção de desvio das frequências alélicas do loco. A heterozigiosidade, o índice de conteúdo polimórfico (PIC) e o Qui-quadrado (χ^2), são alguns dos mais utilizados. Estes parâmetros permitem determinar também, se o marcador molecular utilizado é informativo ou não para a população analisada (SLAVOV *et al.*, 2004).

A heterozigiosidade é uma medida simples da variação genética de uma população em um determinado loco, refletindo a frequência de heterozigotos para esse loco. Considera-se um loco como marcador polimórfico quando a sua heterozigiosidade é ≥ 0.1 e como altamente polimórfico quando a sua heterozigiosidade é ≥ 0.7 (BALZARINI *et al.*, 2006).

O índice de conteúdo polimórfico (PIC) mede o nível de polimorfismo do marcador SSR. Um valor PIC alto indica maior conteúdo de informação de ligamento. Quando o número de alelos é grande, o PIC se aproxima à heterozigiosidade e quando todos os alelos apresentam igual frequência, o PIC é menor que a heterozigiosidade (BALZARINI *et al.*, 2006).

O qui-quadrado (χ^2) permite determinar se a herança dos alelos em uma população determinada está de acordo com a lei de segregação Mendeliana na proporção (1:2:1). Valores de Qui-Quadrado >0.95 , aceitam a hipótese de não independência entre as frequências genóticas observadas e as frequências genóticas esperadas, indicando que as frequências genóticas observadas e esperadas são dependentes (BALZARINI *et al.*, 2006).

Adicionalmente, após verificação dos parâmetros genéticos, neste tipo de estudos de associação entre genótipo e fenótipo, é necessário a utilização de outros tipos de análises estatísticas. A análise de variância (ANOVA), é uma ferramenta estatística que possui várias aplicações comuns no contexto de melhoramento de plantas usada para testar diferenças entre médias de tratamentos quando a variável independente é categórica (por exemplo, os genótipos AA, Aa, aa) e a variável dependente é contínua (nível de tolerância ao déficit hídrico), permitindo avaliar se existe diferença em uma variável contínua entre grupos de tratamento de um experimento (FRANCIS; HEATHER; ROBBINS, 2019).

O mapeamento e identificação de locos de caracteres quantitativos com o uso de marcadores moleculares possibilitam o estudo de características complexas como a tolerância ao déficit hídrico e tem como princípio básico a obtenção de dados fenotípicos em uma progênie e a identificação de regiões no genoma onde o genótipo do marcador pode ser associado ao fenótipo dos indivíduos (BROMAN; SAUNAK, 2009).

Os estudos focados na validação de associações marcador-QTL são desenvolvidos com a intenção de chegar até genes específicos que estejam associados com a característica quantitativa de interesse. Para tal finalidade, é necessária a exploração das regiões associadas para a busca de genes associados com a tolerância ao déficit hídrico. O termo *in silico* é usado para descrever experimentos ou processos conduzidos em um computador ou por meio de simulação computacional, os quais são particularmente úteis no estudo de sistemas complexos ou quando a condução de experimentos pode ser impraticável, demorada ou cara (MORETTI, 2011). O progresso das tecnologias de sequenciamento em combinação com o progresso da bioinformática tem desempenhado um papel na descoberta de genes associados a respostas como a tolerância ao déficit hídrico, fornecendo conhecimentos sobre a base genética de características e alvos potenciais para a seleção de materiais de interesse (PARISSEAU; BERNARDO, 2004). Várias estratégias *in silico* podem ser utilizadas no sentido de identificar, caracterizar e comparar o repertório de genes que participam das principais respostas aos estresses abióticos e bióticos nas plantas. Juntas estas informações devem auxiliar os programas de melhoramento genético da cultura, seja através de métodos convencionais ou da

bioengenharia. De fato, considerando que os métodos de cultivo adotados, na maioria das vezes utilizam pouca tecnologia, reduzem a produtividade e a qualidade de grãos, o uso de cultivares com características genéticas competitivas e eficientes deverão contribuir para alcançar um padrão de agricultura mais sustentável e com maior produtividade (SOARES, 2012). No caso do milho, o projeto MaizeGDB (<https://www.maizegdb.org/>) além de possuir o maior registro da genética da cultura, fornece ferramentas de mineração de dados que permitem fazer pesquisas *in silico* com diferentes objetivos (HARPER *et al.*, 2016).

O objetivo nesta pesquisa foi avaliar associações entre a característica de tolerância ao déficit hídrico em plântulas de milho e marcadores SSR em populações F₂ para serem usados como marcadores de seleção para esta característica, em programas de melhoramento genético de milho. Também foi objetivado o levantamento de genes associados à tolerância ao déficit hídrico, baseado na exploração *in silico* de regiões genômicas relacionadas com esta característica.

2. MATERIAL E MÉTODOS

2.1 OBTENÇÃO DAS POPULAÇÕES SEGREGANTES (FENOTIPAGEM)

Linhagens de milho contrastantes para a característica de tolerância ao déficit hídrico, desenvolvidas no programa de Melhoramento de milho da UFLA, foram usadas para a produção de progênies F_{2:3}. Os cruzamentos dos parentais Linhagem 1 (P1) classificada como tolerante e a linhagem 2 (P2) como suscetível, foram conduzidos por Von Pinho, 2021 na Universidade Federal de Lavras (UFLA), em Lavras, MG (latitude 21°14'S, longitude 40°17'W e altitude de 918.80m). A população F_{2:3} segregante obtida, consistiu de 203 progênies as quais foram classificadas fenotipicamente quanto a tolerância ao déficit hídrico, no estágio de plântulas. A simulação das condições de deficiência hídrica foi feita por meio da utilização de soluções de polietilenoglicol (PEG 6000) com potencial osmótico de -0.6Mpa, para umedecer os papéis de germinação, quando da semeadura. Foram utilizadas 30 sementes de cada progênie, em duas repetições de 15 cada. Após a semeadura, os rolos de papel contendo as sementes, foram acondicionados em sacos plásticos, para evitar a evaporação da água e da

solução de PEG e na sequência mantidos em câmara de germinação do tipo BOD sob temperatura constante de 30°C.

As avaliações foram realizadas aos cinco dias após a instalação do experimento, foram contabilizadas as plântulas que possuíam pelo menos 3cm de raiz principal e pelo menos 3 raízes secundárias. Foram extraídos os valores médios do comprimento da raiz principal e das raízes secundárias.

Os dados foram interpretados estatisticamente por meio da análise de variância, e as médias comparadas pelo teste de Scott-Knott, ao nível de 5%, com o auxílio do programa R[®] (R DEVELOPMENT CORE TEAM, 2019). A característica fenotípica cumpriu com a pressuposição básica de apresentar distribuição normal (Figura 1).

Com base nessa avaliação as progênies foram classificadas quanto à tolerância ao déficit hídrico de acordo com os seguintes critérios: Tolerantes (80% a 100%), Intermediárias (79% a 50%) e Suscetíveis (abaixo ou igual a 49%), com base nos resultados observados aos 5 dias. Essa classificação foi realizada de acordo com a pesquisa desenvolvida por (SALGADO *et al.*, 2008), para a característica de tolerância de sementes de milho às altas temperaturas de secagem. Foi observada alta eficiência desta metodologia para a seleção de genótipos para a tolerância ao déficit hídrico, Figura 2.

Figura 1- Segregação da característica fenotípica na população F_{2:3}.

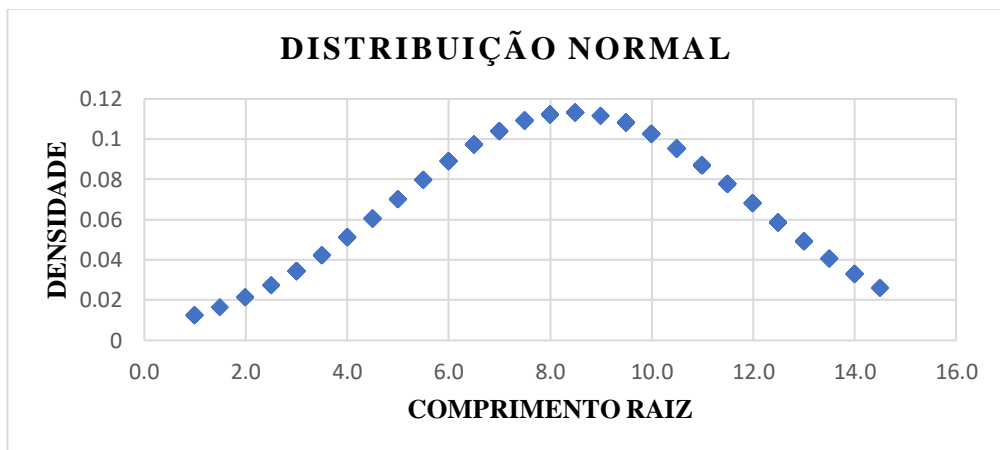
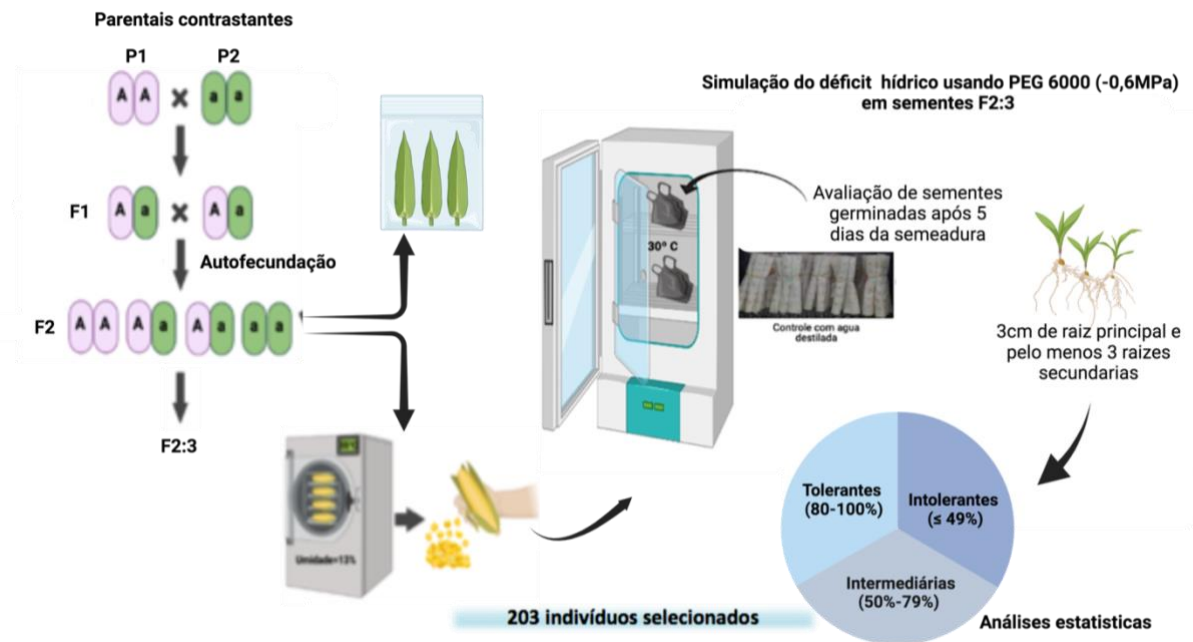


Figura 2- Esquema ilustrativo da metodologia de fenotipagem e obtenção da população segregante.



2.2 GENOTIPAGEM DA POPULAÇÃO F₂ SEGREGANTE

2.2.1. EXTRAÇÃO DE DNA E SELEÇÃO DOS MARCADORES SSR

Durante a fase de experimentos de campo, foram amostradas folhas das progênies e mantidas a -80°C até as análises moleculares. As análises moleculares foram realizadas no Laboratório Central de Pesquisa em Sementes (LCPS) e no Laboratório Central de Biologia Molecular (LCBM) da Universidade Federal de Lavras.

O DNA genômico foi extraído a partir dos parentais P₁ e P₂ contrastantes para a característica de tolerância ao déficit hídrico e das 203 plantas das progênies usando o método CTAB 2X modificado por (VILLANUEVA-MEJÍA *et al.*, 2015). Cada amostra foi analisada por eletroforese em gel de agarose 1% e a concentração de DNA foi quantificada usando o NanoDrop 1000 (Thermo Scientific, Wilmington, DE, USA).

Para a seleção dos marcadores SSR (*Simple Sequence Repeats*), foi realizada uma revisão detalhada a partir da base de dados Maize GBD (<https://www.maizegdb.org/>) (HARPER *et al.*,

2016) dos microssatélites mapeados em cada um dos 10 cromossomos de milho com alguma relação às respostas fisiológicas associadas a tolerância ao déficit hídrico e bem distribuídos no genoma. Foram selecionados mínimo 10 marcadores por cromossomo, totalizando 105 microssatélites analisados (Tabela S1, material anexo).

A padronização das condições de PCR (*Polimerase Chain Reaction*) para a amplificação dos 105 microssatélites, foi feita por PCR em gradiente de temperatura de anelamento (T_a) entre 50°C e 70°C. Cada reação de PCR consistiu de um volume final de 15 μ l, contendo 50 ng de DNA, com 1.5 μ l de tampão de reação 10X, 0.6 mM de MgCl₂, 0.3 μ M de cada dNTP, 1 U de Taq-DNA Polimerase (Invitrogen), e 0.3 μ M de *primers Forward* e *Reverse* específicos e água. Para as amplificações foi usado o termociclador T100 Thermal Cycler (Bio-Rad). O protocolo de amplificação usado foi: desnaturação inicial a 94°C por 5 min, seguido de 35 ciclos de desnaturação a 94°C por 30 segundos, anelamento dos *primers* a 55°C-68°C por 1 minuto, extensão a 72°C por 1 min, seguido de uma extensão final a 72°C por 7 min e finalização da reação a 12°C por tempo infinito (∞). Os resultados das amplificações foram visualizados por meio de eletroforese em gel de agarose 2%, 80 V e 60 min de corrida. Para definir o tamanho dos fragmentos, foi utilizado o marcador de tamanho molecular 50 pb DNA Ladder (Invitrogen). Posteriormente foi visualizado o resultado usando o transiluminador ultravioleta Spectroline® e o fotodocumentador EDAS 290 Kodak®.

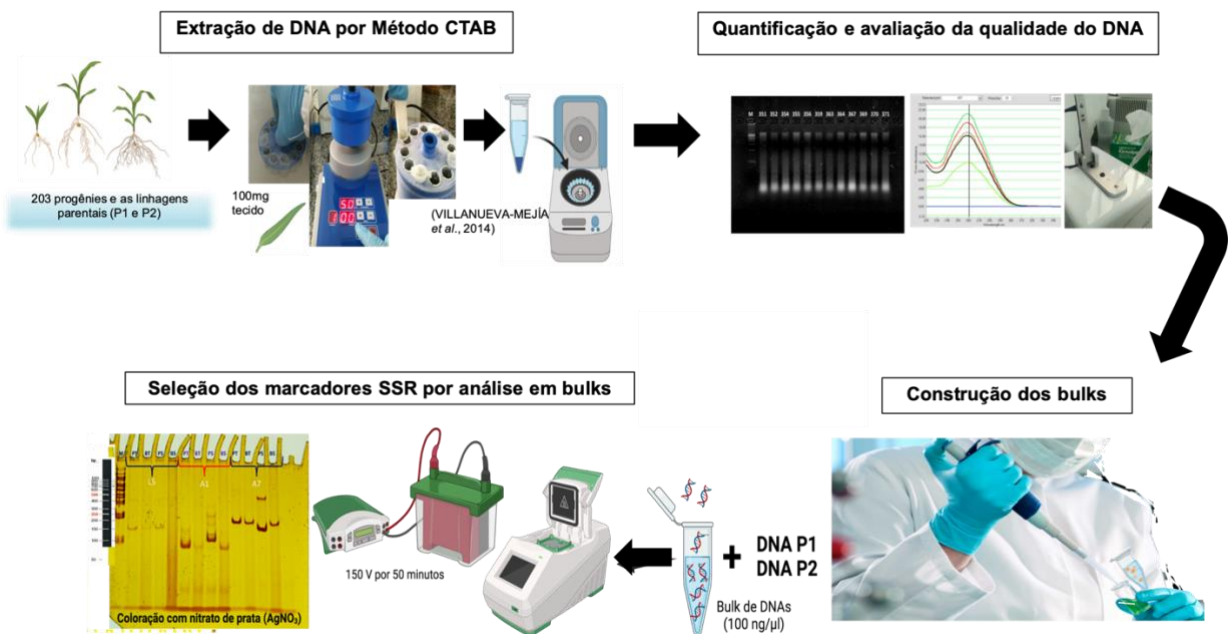
Uma vez estabelecidas as temperaturas de anelamento para os *primers*, as sequências microssatélites SSR foram analisadas de acordo com a estratégia de análise de bulks segregantes baseada na mistura em iguais quantidades de DNA genômico extraído de plantas individuais com o mesmo fenótipo (EMBRAPA CLIMA TEMPERADO, 2007). Foram analisados por PCR dois bulks: um bulk formado por 10 DNAs (100 ng/ μ l) de progênies identificadas como “tolerantes” e um bulk de 10 DNAs (100 ng/ μ l) de progênies identificadas como “suscetíveis” junto com os DNA dos parentais P₁ e P₂.

Os produtos das amplificações foram separados em gel de poliacrilamida 10%, 150 V por 50 minutos. Para a revelação dos géis utilizou-se o método de coloração com nitrato de prata. Após a eletroforese, as placas foram separadas e cada gel foi imerso em 250 ml de solução fixadora (etanol 10%-ácido acético 0.5%) e mantido sob lenta agitação por 20 minutos. Em seguida, cada gel foi imerso em 250 ml de solução de nitrato de prata (AgNO₃) (0.2%) sob agitação lenta por 20 minutos. Os géis, então, foram lavados com água corrente entre cada passo. Os géis foram mantidos imersos sob agitação lenta em aproximadamente 180 ml da solução de revelação (NaOH 3%, formaldeído 0.5%) até a completa visualização das bandas.

Após o aparecimento das bandas os géis foram colocados novamente na solução fixadora. As imagens foram captadas usando o transiluminador de luz branca Biosystems e a câmera Sony Alpha SLT-A57K. O processo de extração e seleção dos marcadores SSR por análises em bulks é resumido na Figura 3.

Após essa avaliação inicial, os *primers* que apresentaram polimorfismo entre as linhagens e os bulks foram utilizados para uma segunda análise de polimorfismo por eletroforese capilar.

Figura 3- Diagrama ilustrativo da metodologia usada para a extração do DNA e as análises em bulks.



2.2.2. GENOTIPAGEM DA POPULAÇÃO F₂ POR ELETROFORESE CAPILAR

Os marcadores SSR previamente selecionados nas análises de bulks em poliacrilamida, foram analisados pelo sistema de eletroforese capilar TapeStation 4150 (Agilent Technologies) usando os kits de análise D1000 DNA screenTape na Universidade EAFIT, Medellín, Colômbia.

As amostras foram preparadas seguindo o protocolo recomendado pelo fabricante Agilent D1000 ScreenTape System Quick Guide, e analisadas usando o *analyzer software*.

Os marcadores mmc0371 do cromossomo 4, bnlg1208 e umc1056 do cromossomo 5, umc1401 do cromossomo 7 e phi027 e phi061 do cromossomo 9 foram os marcadores analisados para a genotipagem das populações segregantes. Foi construída uma matriz com a presença dos alelos para as análises de diversidade genética.

Para a genotipagem utilizando a eletroforese capilar, foi feito um exame cuidadoso de cada eletroferograma gerado pelo programa do sistema de eletroforese capilar *TapeStation* 4150 (Agilent Technologies). O eletroferograma é a representação gráfica dos resultados da eletroforese capilar que ajuda na identificação de alelos falsos ao identificar presença de picos inesperados ou irregularidades nas alturas dos picos reais. Adicionalmente, realizou-se a comparação com os padrões conhecidos, no caso os parentais P1 e P2, para identificar quaisquer discrepância ou inconsistências (TAVARES, 1997).

Para a identificação correta dos alelos próximos, avaliou-se a concentração calibrada [ng/μl], a molaridade do pico [nmol/l] e a porcentagem da área integrada do pico. Quanto maiores forem esses parâmetros, mais confiável é a seleção do pico. Se dois picos estiverem muito próximos e o segundo pico não chegasse pelo menos ao 30% do sinal do pico mais forte, o pico foi considerado “falso”. Foram selecionados então segundos picos com 50% ou mais do sinal do pico mais forte, isto foi necessário para a genotipagem de quase todos os indivíduos heterozigotos.

2.3. VARIABILIDADE GENÉTICA DOS MARCADORES SSR

Foram identificados 11 *primers* dos 14 pré-selecionados como polimórficos, no entanto, foram escolhidos 6 deles para a genotipagem das populações F₂ segregantes completas para a característica de tolerância ao déficit hídrico usando o sistema de eletroforese capilar *TapeStation* 4150 (Agilent Technologies).

Os dados de genotipagem dos locos nas progênes foram analisados no software *Cervus* (KALINOWSKI; TAPER; MARSHALL, 2007; MARSHALL *et al.*, 1998). Foi determinada a quantidade de alelos nas progênes, os índices de conteúdo polimórfico de cada loco, as frequências alélicas e genotípicas, heterozigosidades observadas e a segregação do loco foi verificada quanto ao desvio da proporção mendeliana normal pelo teste padrão do Qui-quadrado (χ^2).

A heterozigosidade foi estimada segundo a fórmula:

$$H^{\wedge} = \frac{N}{N - 1} \left(1 - \sum_{i=1}^l P_i^{\wedge 2} \right)$$

Em que:

N: tamanho da amostra;

$\sum_{i=1}^l P_i^2$: frequência esperada para l genótipos homocigotos.

Para a estimar o polimorfismo, também é comum usar o índice de conteúdo polimórfico (PIC). Quando o número de alelos é grande, o PIC se aproxima à heterozigosidade e quando todos os alelos apresentam igual frequência, o PIC é menor que a heterozigosidade (BALZARINI *et al.*, 2006). O índice de conteúdo polimórfico para cada marcador foi estimado segundo a fórmula:

$$PIC = 1 - \sum_{i=1}^l P_i^2 - 2 \sum_{i=2}^l \sum_{j=1}^{i-1} (P_i^2 P_j^2)$$

$$PIC = 2 \sum_{i=2}^l \sum_{j=1}^{i-1} (P_i P_j) - 2 \sum_{i=2}^l \sum_{j=1}^{i-1} (P_i^2 P_j^2)$$

$$PIC = 2 \sum_{i=2}^l \sum_{j=1}^{i-1} [P_i P_j (1 - P_i P_j)]$$

O teste do Qui-cuadrado (χ^2) avalia a probabilidade de que a distribuição ou herança dos alelos, tenha acontecido seguindo o padrão de segregação Mendeliana (BALZARINI; DI RIENZO, 2003). O teste do χ^2 foi realizado baseado na diferença entre os valores de heterozigosidade observada e esperada, segundo a fórmula:

$$\chi^2 = \sum \left(\frac{H^{\wedge} \text{observada} - H^{\wedge} \text{esperada}^2}{H^{\wedge} \text{esperada}} \right)$$

2.4. ANÁLISES DE ASSOCIAÇÃO ENTRE OS SSR POLIMÓRFICOS E A TOLERÂNCIA AO DÉFICIT HÍDRICO (ASSOCIAÇÃO MARCADOR-QTL)

Os dados da genotipagem foram transformados a dados binários; presença ou ausência dos alelos, e dados fenotípicos (porcentagens de tolerância ao déficit hídrico). Posteriormente, foram analisados por meio de uma análise de variância (ANOVA) com o auxílio do software *InfoGen* (BALZARINI; DI RIENZO, 2003), usando como variável de classificação o nível de tolerância ao déficit hídrico e seguindo o modelo estatístico:

$$\gamma_{ij} = \mu + p_i + e_{ij}$$

Em que:

γ_{ij} : observação do genótipo i na população j;

μ : constante geral associada a todas as observações;

p_i : efeito fixo do genótipo i;

e_{ij} : erro experimental aleatório associado à observação, γ_{ij} .

O efeito médio da substituição alélica foi determinado codificando os genótipos como 0 (suscetível homozigoto), 1 (heterozigoto) e 2 (tolerante homozigoto) para representar os dois alelos de cada locus, utilizando o seguinte modelo de regressão linear:

$$\gamma = \alpha + \beta x$$

Em que:

α : Constante;

βx : coeficiente de regressão.

Os dados também foram analisados por meio de uma análise de contingência para a determinação da associação do loco com a característica quantitativa de tolerância ao déficit hídrico. Ambas as análises foram feitas usando o programa *InfoGen*.

2.5. IDENTIFICAÇÃO *IN SILICO* DE GENES ASSOCIADOS AOS LOCOS SSR **bnlg1208 E phi027**

Foi feita uma pesquisa *in silico* para 100 Kb *downstream* e *upstream* dos locos bnlg1208 e phi027, com a finalidade de encontrar outros genes próximos a essas regiões e a sua função associada à tolerância ao déficit hídrico. disponíveis

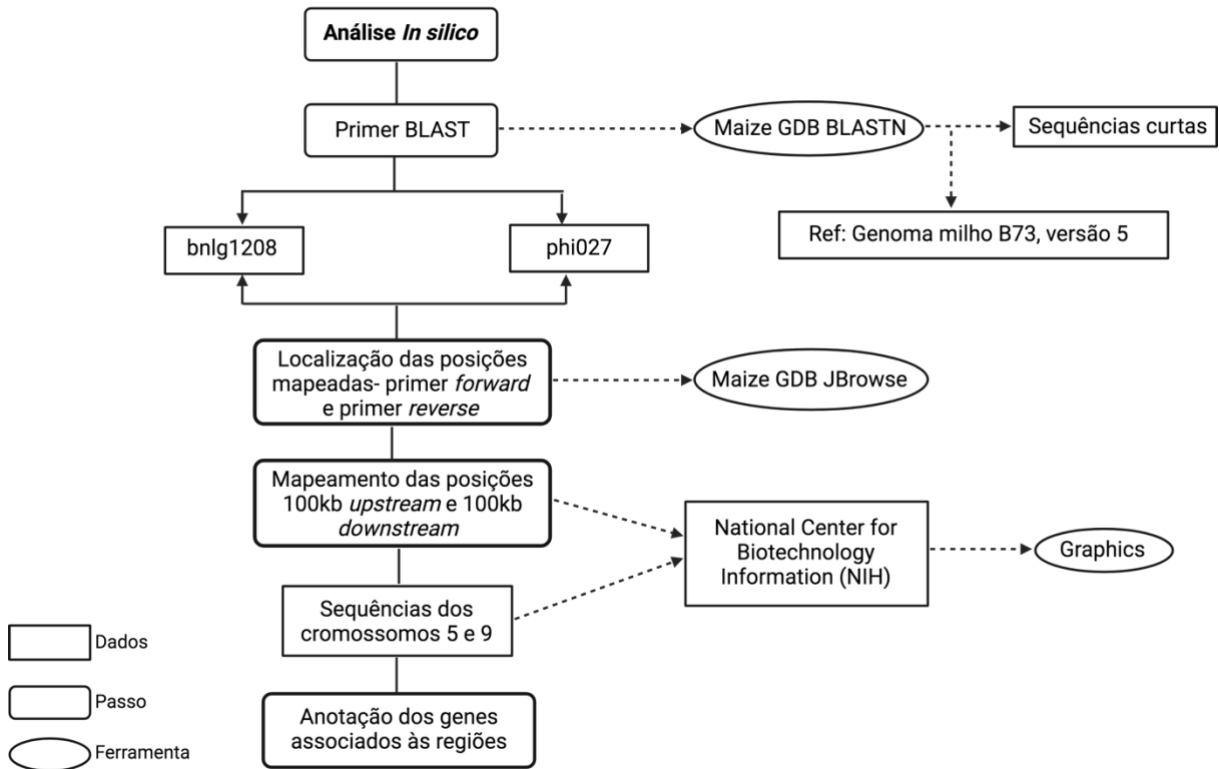
Para isto foi usada a ferramenta BLAST do MaizeGBD (https://www.maizegdb.org/popcorn/search/sequence_search/home.php?a=BLAST_UI) para mapear cada par de *primers* dos marcadores bnlg1208 e phi027. Para o BLAST dos *primers* foi usado como genoma de referência o genoma B73, versão 5. Os parâmetros de busca foram otimizados para sequências curtas (*Short sequences*) (Figura S1, material anexo).

O resultado de cada BLAST foi visualizado na ferramenta JBrowse do MaaizeGBD (<https://jbrowse.maizegdb.org/>) junto com as anotações para essa região nessa base de dados.

Posteriormente, as posições exatas localizadas 100kb *upstream* e *downstream* da localização dos locos bnlg1208 (153,588,898..153,789,017) e phi027 (25,021,113..25,241,244), foram visualizadas no sitio web do *National Center for Biotechnology Information* (NIH) (<https://www.ncbi.nlm.nih.gov/>). Para isto foram usadas as sequências dos cromossomos 5 e 9 do genoma B73 (*accession numbers*: CM000781.4 e CM000785.4, respectivamente) e a ferramenta de visualização *graph* do (<https://www.ncbi.nlm.nih.gov/>). Uma vez localizadas as posições, foi feita uma busca manual dos genes anotados dentro de cada região para os locos estudados. Os passos para a análise *in silico* são descritos de forma simplificada na Figura 4.

Os genes identificados próximos aos locos de interesse, foram investigados por meio de dados disponíveis na literatura e descreveu-se as funções preditas dos genes candidatos associados á tolerância ao déficit hídrico.

Figura 4- *Framework* computacional simplificado baseado nas análises bioinformáticas que foram empregadas para a identificação de genes associados aos locos relacionados com a tolerância ao déficit hídrico em Milho.



3. RESULTADOS E DISCUSSÃO

A padronização das temperaturas de anelamento dos SSR permitiu fazer as análises de polimorfismo em bulks obtendo amplificação em 96 SSR dos 105 testados. Uma das causas da não amplificação de um produto de PCR pode ser a temperatura de anelamento. A redução da temperatura de anelamento durante a PCR pode reduzir a especificidade de ligação do *primer* e pode resultar na produção de múltiplos produtos inespecíficos. No caso contrário, o aumento da temperatura de anelamento, pode aumentar a especificidade de ligação do *primer*, porém, se a temperatura for demasiadamente elevada em relação à temperatura de *melting* (T_m) do *primer*, a reação pode falhar. A faixa recomendada para temperaturas de fusão na PCR é de 48 a 68°C, assim como foi testado no presente estudo. É importante otimizar e manter a temperatura de anelamento para garantir uma amplificação específica e eficiente das sequências SSR alvo e

para obter resultados confiáveis e precisos na genotipagem (MASON, 2015). A temperatura de anelamento para genotipagem de SSR está normalmente em torno de 55°C, não obstante, a temperatura de anelamento ótima, é influenciada pela temperatura de *Melting*, que a sua vez é determinada por fatores como comprimento do *primer*, conteúdo de GC e composição da sequência (LIANG *et al.*, 2023; NAEEM; MIRZA, 2018).

Os processos de genotipagem e fenotipagem de uma característica de interesse como a tolerância ao déficit hídrico, precisam de populações grandes, o que pode ser um processo trabalhoso e relativamente caro em termos de consumíveis. A técnica BSA (análise em bulks segregantes) se torna interessante para a categorização de populações com base na expressão de uma característica específica segregante e permite simplificar o processo de genotipagem em poucas amostras com fenótipo comum (NAEEM; MIRZA, 2018; QUARRIE *et al.*, 1999).

Neste estudo, 14 *primers* foram polimórficos entre os bulks, exibindo 14.73% de polimorfismo nas análises em géis de poliacrilamida (Figura 5). Esses 14 marcadores SSR polimórficos detectados, se encontraram distribuídos em 8 dos 10 cromossomos do milho (Tabela 1).

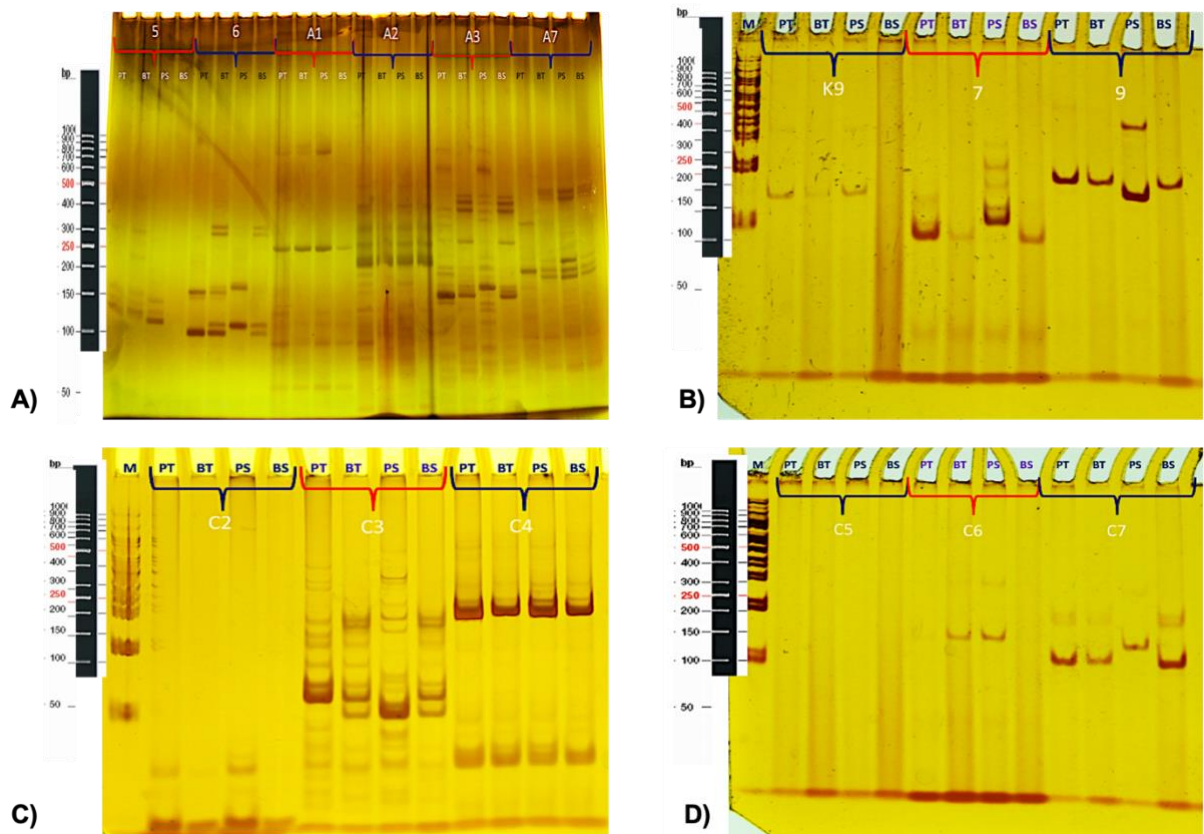
Tabela 1- Marcadores SSR polimórficos nas análises em bulks por eletroforese em gel de poliacrilamida.

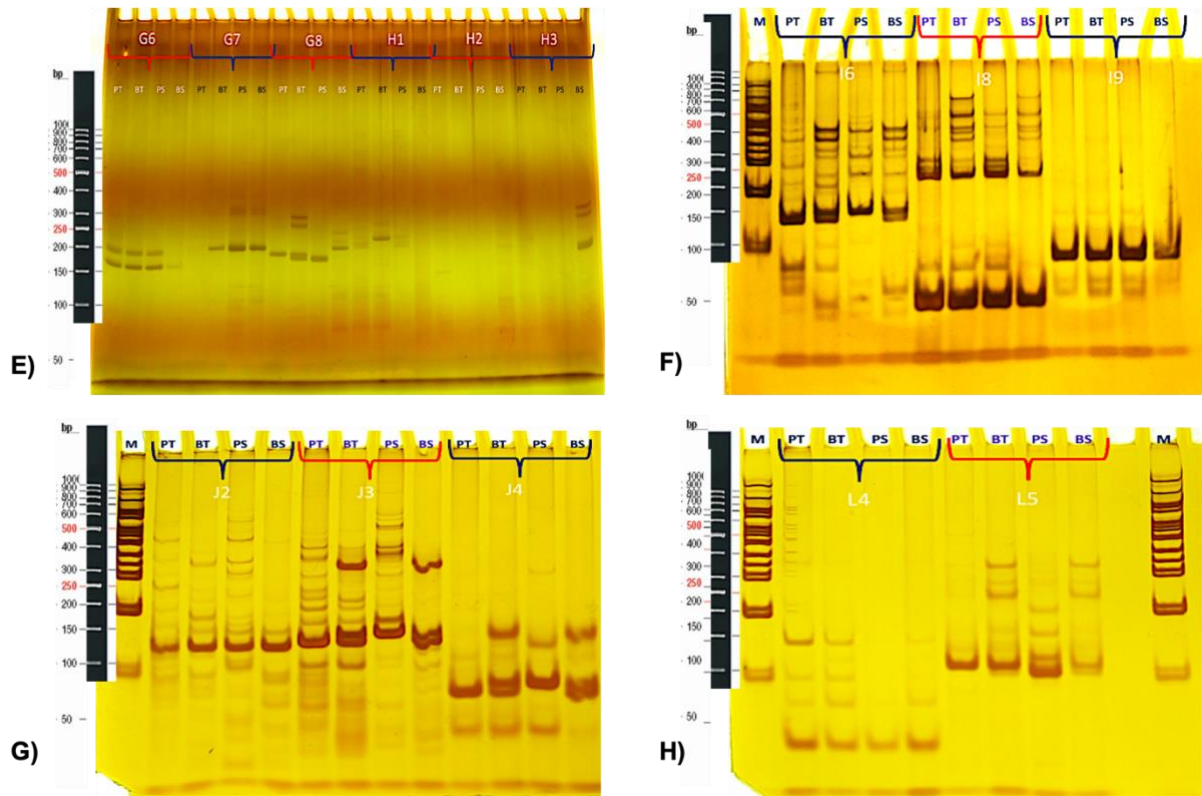
SSR Maize GBD	Sequência (5' - 3')	Repetição	T _{ann} (°C)	Tamanho esperado	Cromossomo
umc1448-F	ATCCTCTCATCTTTAGGTCCACCG	(GCT) ₅	68	137-161	2
umc1448-R	CATATACAGTCTCTTCTGGCTGCTCA				
umc1080-F	GAGGAGAAAAGGAGATGGAAAAGC	(GA) ₁₅	65	89-123	2
umc1080-R	AGATGCCGCGAGAAGATTCTAAACA				
phi053-F	CTGCCTCTCAGATTCAGAGATTGAC	ATAC	65	154-200	3
phi053-R	AACCCAACGTACTIONCCGGCAG				
phi047-F	GGAGATGCTCGCACTGTTCTC	ATC	68	-	3
phi047-R	CTCCACCCTCTTTGACATGGTATG				
bnlg1601-F	CAGACCAGAGACCATCTGCA	AG ₍₂₄₎	68	-	3
bnlg1601-R	ATCGTGCGCTAGTCCAGAGT				
mmc0371-F	CCACGCACCTCTTGTAAAC	(GA) ₄ N ₁ (GA) ₂₃ (GGA) ₁	65	186-200	4
mmc0371-R	GGGAACCTACAGCTTGGT	(GAA) ₂			
bnlg1208-F	GCTGTGATGGTGAGACGAGA	AG ₍₁₀₎	68	-	5
bnlg1208-R	GCAGGCACTACTAAAACCGC				
umc1056-F	CGGATCGCTTTTTACCGTCTATAA	(AGCA) ₄	65	105-161	5
umc1056-R	AGCAAGAGTAGCGTTCCATTTCAG				

umc1401-F	CTCTGGTCCATCCTCATCGACT	(CCA) ₄	65	151	7
umc1401-R	TCTCTTGATCACATATCGATCCCA				
phi080-F	CACCCGATGCAACTTGCCTAGA	AGGAG	68	155	8
phi080-R	TCGTACAGTTCCACGACATCAC				
phi022-F	TGCGCACCAGCGACTGACC	GTGC	68	137-165	9
phi022-R	GCGGGCGACGTTCCAAAC				
phi027-F	CACAGCACGTTGCGGATTTCTCT	GCGCT	68	-	9
phi027-R	GCGTACGTACGACGAAGACAC				
phi061-F	GACGTAAGCCTAGCTCTGCCAT	TTCT-GTAT	68	-	9
phi061-R	AAACAAGAACGGCGGTGCTGATTC				
nc009-F	CGAAAGTCGATCGAGAGACC	(AG) _x	68	-	6
nc009-R	CCTCTTTCACCCCTTCCTT				

Fonte: Da autora (2024)

Figura 5-Resultados de polimorfismo dos marcadores SSR nas análises em bulks em géis de poliacrilamida.





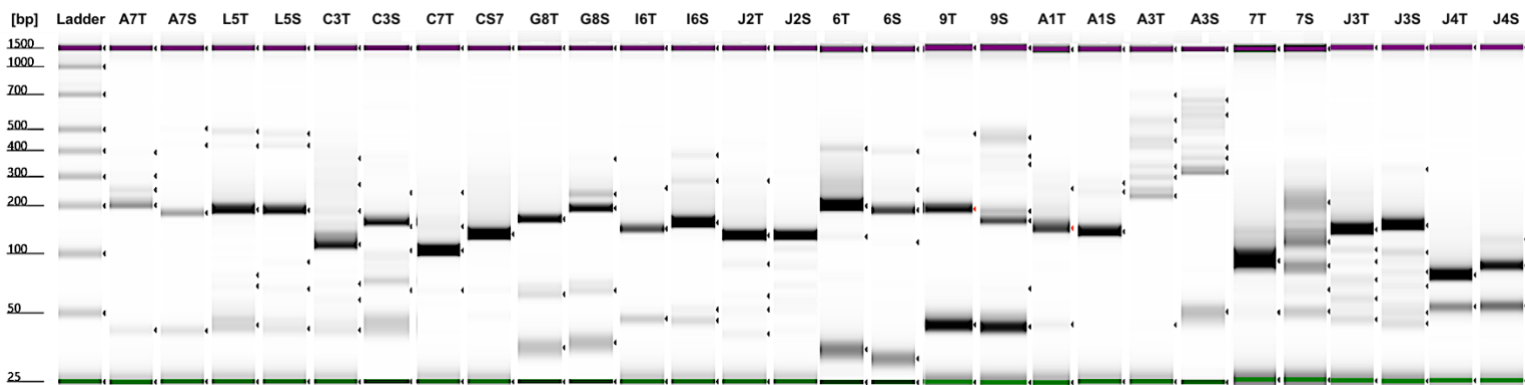
A) *umc1448*, *phi047*, *bnlg1601* e *mmc0371*; B) *umc1080* e *phi053* C) *bnlg1208*; D) *umc1056*; E) *umc1401*; F) *phi080*; G) *phi022*, *phi027*, *phi061*; H) L5. **PT**- Parental tolerante, **BT**- Bulk de indivíduos tolerantes, **PS**- Parental suscetível, **BS**- Bulk de indivíduos suscetíveis, **M**- Marcador de tamanho molecular.

Fonte: Da autora (2024)

Os resultados do polimorfismo dos 14 *primers* selecionados pelas análises em gel de poliacrilamida, foram confirmados usando o sistema de eletroforese capilar, por meio da análise dos parentais.

Pelos resultados das análises dos 14 *primers* por eletroforese capilar foi verificado que os marcadores SSR *nc009* e *phi022* não eram polimórficos como pareciam ser nos resultados obtidos com poliacrilamida. Para o marcador *bnlg1601* houve ampliações inespecíficas que não permitiram fazer a identificação de alelos específicos (Figura 6). Para os marcadores SSR restantes foi possível confirmar o polimorfismo entre os parentais e os bulks. O número total de alelos polimórficos identificados foram 22, e para cada marcador SSR foram evidenciados 2 alelos para os parentais. O tamanho dos produtos de PCR variou de 81 a 205 pb (Tabela 2, Figura 6).

Figura 6- Resultados das análises de polimorfismo entre os parentais P1 (tolerante) e P2 (suscetível), no sistema de eletroforese capilar.



Fonte: Da autora (2024)

Tabela 2- Tamanho e numero de alelos polimórficos determinados pelo sistema de eletroforese capilar.

SSR Maize GBD	Tamanho esperado (pb)	Tamanho (pb)	Numero de alelos
umc1448	137–161	194, 205	2
umc1080	89-123	93, 123	2
phi053	154–200	177, 97	2
phi047	-	148, 156	2
bnlg1601	-	Inespecífico	-
mmc0371	186-250	186, 200	2
bnlg1208	-	96, 120	2
umc1056	105–161	109, 138	2
umc1401	151	130, 147	2
phi080	155	152, 165	2
phi022	135–154	138	1
phi027	-	150, 159	2
phi061	-	81, 89	2
nc009	-	110	1

Fonte: Da autora (2024)

Dos 11 *primers* polimórficos foram escolhidos os marcadores mmc0371, bnlg1208, umc1056, umc1401, phi027 e phi061, para a genotipagem das populações segregantes, por serem polimórficos com marcas facilmente identificáveis. Foi construída uma matriz com a presença dos alelos para as análises de diversidade genética por meio da análise de bandas produzidas por eletroforese capilar de forma cuidadosa sobretudo para pontuar alelos que se encontravam muito próximos (Figura 7), pois é frequente na eletroforese capilar encontrar marcadores que apresentam o fenômeno conhecido como "gagueira" pode causar falsas marcas adicionais

separadas por intervalos de 0,5-2 pb. Essas falsas marcas tendem a se manifestar em padrões específicos e podem ser identificados como tal (MASON, 2015).

Figura 7- Eletroferograma e análise de dados gerados pelo sistema de eletroforese capilar TapeStation 4150 (Agilent Technologies) para o caso de dois alelos muito próximos.



Gel de eletroforese capilar com as bandas. Do lado direito se apresenta o eletroferograma. O círculo destaca dois alelos que apresentam picos muito próximos. A seta indica o tamanho molecular de cada alelo. A tabela destacada no quadrado, amostra os dados de qualidade dos picos de forma condensada.

Fonte: Da autora (2024)

Ressalta-se que as falsas marcas podem ser causadas por vários fatores, incluindo erros técnicos, polimorfismos no local de ligação do *primer* ou mutações de deleção. Portanto, é crucial analisar cuidadosamente os resultados da eletroforese e considerar múltiplos fatores antes de concluir a presença de alelos falsos (MASON, 2015). No presente estudo, foram consideradas marcas próximas devido a que segregaram corretamente na população analisada além de apresentar tamanhos com uma diferença superior a 8pb.

A análise de descrição dos dados dos 6 marcadores SSR que geraram um total de 12 bandas polimórficas, com número médio de alelos efetivos (N_e) de 2 alelos por loco. O número de alelos foi similar ao relatado por Kumari *et al.* (2018), que avaliaram um total de 22 marcadores SSR e só 9 foram polimórficos, a maioria deles apresentaram também dois alelos por loco e no máximo 3. A diferença nos números de alelos obtidos entre os diferentes estudos feitos em milho, pode ser devida ao tamanho das amostras em estudo e às metodologias utilizadas para a

detecção de marcadores polimórficos que influenciam nas diferenças alélicas ou na uniformidade baseada em pedigrees (KUMARI *et al.*, 2018).

Nas análises de diversidade genética foi possível determinar os valores de PIC para cada marcador. Todos os marcadores apresentaram capacidade discriminatória e conteúdo de informação polimórfica semelhante, com valores PIC intermediários, entre 0.365 e 0.375. O valor PIC indica a informatividade dos locos SSR e sua capacidade de detectar diferenças entre os genótipos com base em suas relações genéticas (BOCIANOWSKI *et al.*, 2021).

Para o loco phi061 foi obtido o valor de PIC mais alto (0.375) igual ao resultado obtido por Bocianowski *et al.* (2021) com o mesmo marcador SSR (phi061) e um PIC de 0.497. Já o loco umc1056 foi o menos polimórfico, cujo valor do PIC foi 0.365 (Tabela 3).

O valor médio do PIC foi 0.369, consideravelmente mais alto ao obtido por Kumari *et al.* (2018) com um valor médio do PIC de 0.297 em 9 marcadores SSR e Bocianowski *et al.* (2021) com um PIC médio de 0.274 em 30 marcadores SSR e também comparáveis com os resultados obtidos por Kumar *et al.* (2016) que usaram um conjunto de 22 *primers* em 13 genótipos de milho e encontraram valores de PIC médio de 0.304. Valores de PIC baixos, podem ser obtidos devido à alta frequência de dois alelos, à estreita base genética dos genótipos de milho ou devido ao uso particular de *primers* tri/tetra/penta-nucleotídeos.

Em contrapartida, o valor PIC médio obtido neste trabalho foi inferior comparado com os obtidos por Shimelis *et al.* (2015) com PIC de 0.51 em 30 SSR e por Belalia *et al.* (2019), PIC de 0.57 em 18 SSR. Os valores mais elevados de PIC relatados por outros autores, provavelmente resultaram do uso de metodologias de análise diferentes, ao menor número de linhagens endogâmicas utilizadas e a baixa frequência alélica dos genótipos estudados, indicando a presença de nível moderado de diversidade genética entre as linhagens endogâmicas (ADEBAYO *et al.*, 2015).

Outro parâmetro analisado nos testes de diversidade genética foi a heterozigosidade dos locos. Ela estima a probabilidade de um indivíduo ser heterozigoto para aquele loco na população (H_o). Ela é calculada realizando uma estimativa imparcial com base no valor esperado condicional das frequências alélicas (heterozigosidade esperada ou imparcial de Nei (H_e)) (NEI, 1978).

Por meio dos valores de heterozigosidade observada para os marcadores bnlg1208, umc1056, p-umc1401 e phi027, com valores H_o de 0.390 0.431; 0.385 e 0.385. Isso quer dizer que a aproximadamente a metade da população apresentava genótipos homozigotos e a outra metade alelos heterozigotos (BALZARINI *et al.*, 2006). No entanto, para os marcadores mmc0371 e phi061 com valores H_o de 0.139 e 0.242, indicando que não cumprem com o padrão de segregação mendeliana na população analisada.

Tabela 3- Medidas de diversidade genética por loco.

Loco	repetição	K	N	H_o	H_e	PIC
mmc0371	(GA) ₄ N ₁ (GA) ₂₃ (GGA) ₁ (GAA) ₂	2	187	0.139	0.484	0.366
bnlg1208	AG ₍₁₀₎	2	187	0.390	0.490	0.369
umc1056	(AGCA) ₄	2	204	0.431	0.482	0.365
umc1401	(CCA) ₄	2	205	0.385	0.498	0.373
phi027	GCGCT	2	192	0.385	0.489	0.369
phi061	TTCT-GTAT	2	195	0.242	0.500	0.375

K- numero de alelos, **N**- numero de indivíduos genotipados, **H_o** - heterozigosidade observada, **H_e** - heterozigosidade esperada, **PIC**- Conteúdo de informação Polimórfica.

Fonte: Da autora (2024)

Pelos resultados da análise Qui-Quadrado (χ^2) para o estudo da herança na população F_2 , verifica-se que a segregação populacional estava de acordo com a segregação Mendeliana em uma proporção de 1:2:1 para os marcadores SSR bnlg1208, umc1056, umc1401 e phi027, com valores de χ^2 superiores a 0.95. Valores de Qui-Quadrado >0.95 , aceitam a hipótese de não independência entre os intervalos de frequências genotípicas observadas e as frequências genotípicas esperados, ou seja, as frequências genotípicas observadas e esperadas são dependentes (BALZARINI *et al.*, 2006). Entretanto, para os marcadores mmc0371 e phi061 o padrão de segregação não está de acordo com a razão da lei de segregação mendeliana (1:2:1), com valores Qui-quadrado de 0.74 e 0.87, respectivamente, indicando que as frequências genotípicas observadas são independentes das esperadas. O valor-p indica a probabilidade de que as diferenças entre os valores observados e esperados sejam devidas ao acaso (Tabela 4).

Tabela 4- frequências genotípicas relativas por loco.

Loco	Genótipo	H_o	H_e	Qui-Cuadrado (χ^2)	Valor-p
mmc0371	186-186	0.34	0.25	0.74	0.106581
	186-200	0.14	0.5		
	200-200	0.52	0.25		
bnlg1208	96-96	0.23	0.25	0.95	0.003240**
	96-120	0.39	0.5		
	120-120	0.38	0.25		
umc1056	109-109	0.38	0.25	0.95	0.003285**
	109-138	0.42	0.5		
	138-138	0.20	0.25		
umc1401	130-130	0.35	0.25	0.97	0.001595**
	130-147	0.39	0.5		
	147-147	0.27	0.25		
phi027	150-150	0.39	0.25	0.95	0.003840**
	150-159	0.39	0.5		
	159-159	0.23	0.25		
phi061	81-81	0.36	0.25	0.87	0.025263
	81-89	0.24	0.5		
	89-89	0.40	0.25		

(**) indica um valor-p significativo em $\alpha = 1\%$

Fonte: Da autora (2024)

Os quatro marcadores SSR que cumpriram com a lei de segregação de alelos Mendeliana em uma proporção de 1:2:1 foram analisados por meio de três análises estatísticas: 1) análise da variância; 2) análise de regressão linear; 3) análise de contingência.

O coeficiente de correlação (r) quantifica a relação linear entre duas variáveis e varia de -1 a 1. Um valor de 0 indica que não há associação entre as duas variáveis, um valor maior que 0 indica uma associação linear positiva (associação direta) e um valor negativo indica uma relação linear negativa (associação inversa). Por outro lado, o coeficiente de determinação (R^2), é o quadrado do coeficiente de correlação (r), que quantifica a proporção da variância na variável dependente que pode ser explicada pela variável independente. Em outras palavras o R^2 explica a porcentagem de variação fenotípica influenciada por um marcador genético determinado (MIKIĆ *et al.*, 2020; SCHOBER; SCHWARTE, 2018).

O resultado da análise da variância, permitiu observar que os locos bnlg1208 e phi027 estão estatisticamente associados com variável de tolerância ao déficit hídrico, apresentando coeficientes de determinação de 0.4 e 0.19, respectivamente (Tabela 5). Os valores de significância (valores-p) para esses locos também foram altamente significativos ($p < 0.0001$).

Tabela 5- Análises da variância entre os genótipos por loco.

Loco	N	R²	R²(%)	Valor-p
bnlg1208	187	0.4	40.0	<0.0001
umc1056	204	4.60E-03	0.046	0.6293
umc1401	205	4.80E-03	0.048	0.6148
phi027	192	0.19	19.0	<0.0001

N- numero de observações, **R²**- coeficiente de determinação, **R² (%)** - porcentagem da variância fenotípica explicada por cada marcador.

Fonte: Da autora (2024)

Os coeficientes de determinação obtidos neste estudo foram interpretados como a estimativa da proporção da variância fenotípica explicada por cada marcador. Os marcadores umc1056 e umc1401 não explicaram nenhuma variância fenotípica na população estudada. Pelo caso contrario, os marcadores bnlg1208 e phi027 apresentaram altas porcentagens de variância fenotípica explicada pelos marcadores com 40% e 18%, respectivamente.

Xiao *et al.* (2021) estudaram o estresse hídrico em milho usando 121 marcadores SSR e associação às características fenotípicas de rendimento de grãos e componentes de rendimento. Eles encontraram 3 regiões associadas ao estresse hídrico. A primeira localizada no cromossomo 1 associada ao marcador bnlg176, as outras duas regiões associadas ao estresse hídrico foram localizadas no cromossomo 9 perto dos marcadores umc1657 e bnlg1525. Os valores de R² (%) para os marcadores bnlg176 umc1657 e bnlg1525 estiveram nas faixas de 5.25-7.42; 6.82-13.69 e 4.21-31.28, respectivamente. Esse resultado deixa em evidencia que foram significativamente menores aos encontrados nos locos bnlg1208 e phi027 do presente trabalho.

Mikić *et al.* (2020) testaram a tolerância ao déficit hídrico de linhagens de milho de elite usando microssatélites SSR, para as características de rendimento de grãos e intervalo entre florescimentos feminino e masculino (ASI). A análise de associação indicou correlações significativas entre ASI e rendimento de grãos e três microssatélites nos cromossomos 3, 6 e 9 (bnlg1525, bnlg238 e umc1025, respectivamente). No entanto, eles encontraram porcentagens de variância fenotípica menores para ambas características analisadas, entre 7.0 e 8.2% para o rendimento de grãos e entre 10.0% a 22.0% para a característica ASI.

(SALGADO, 2005) encontrou três marcadores SSR associados significativamente com a característica de tolerância à alta temperatura de secagem em sementes de milho. Os locos nc0013, umc2059 e phi014 apresentaram valores de R^2 (%) menores aos reportados neste estudo (10.61; 10.05 e 9.70, respectivamente) porém, também foram associações significativas entre as características analisadas.

As comparações entre médias por loco evidenciaram que os locos bnlgl208 e phi027 segregam com significância estatística os diferentes níveis de tolerância ao déficit hídrico dos genótipos avaliados. O genótipo 2 foi usado como condição para caracterizar as progênies homozigotas tolerantes, o genótipo 1 para indicar as progênies heterozigotas e o genótipo 0 para indicar as progênies homozigotas suscetíveis (Tabela 6; Figura 8). Para ambos locos o genótipo suscetível (0) foi o que apresentou médias menores e o genótipo tolerante (2) apresentou médias maiores em quanto a porcentagem de tolerância ao déficit hídrico.

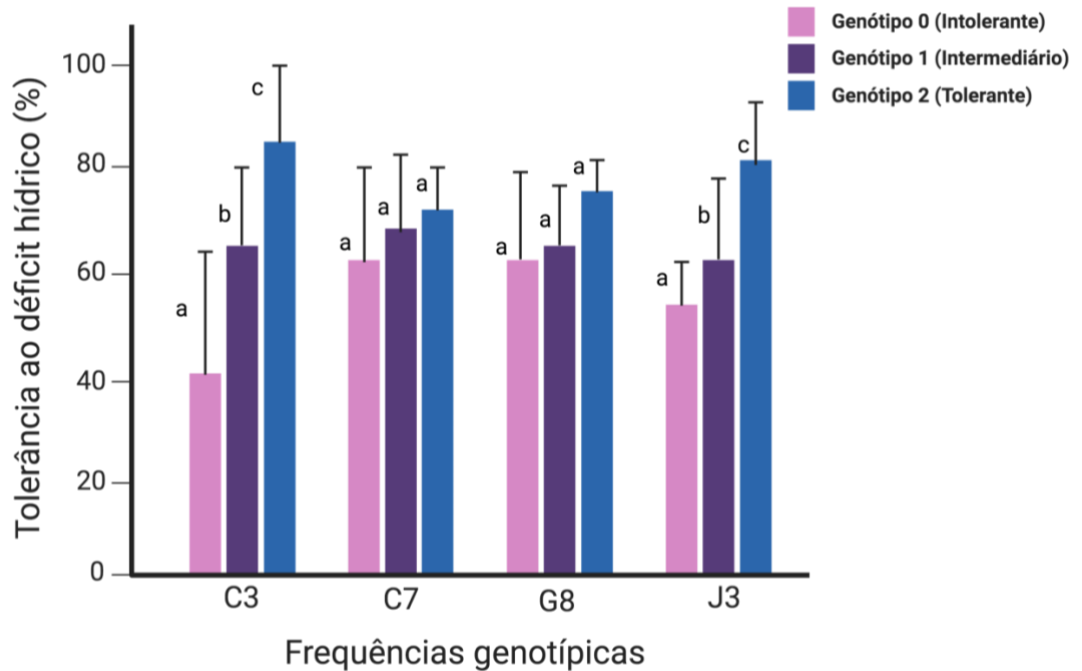
Tabela 6- Dados estatísticos da comparação entre médias resultado da ANOVA.

Loco	Genótipos	(%) Tolerância	n	E.E
bnlg1208	2	81.76	74	2.95 a
	1	66.09	72	2.23 b
	0	41.17	41	2.24 c
umc1056	2	68.80	41	2.62 a
	1	67.85	78	2.73 a
	0	64.95	85	3.77 a
umc1401	2	69.68	55	2.87 a
	1	66.32	79	2.72 a
	0	65.61	71	3.26 a
phi027	2	79.60	74	3.29 a
	1	62.70	74	2.54 b
	0	54.01	44	2.54 c

E.E- erro padrão. Médias com uma letra comum não são significativamente diferentes ($p > 0.05$)

Fonte: Da autora (2024)

Figura 8- Comparação de médias entre os genótipos por loco (ANOVA).



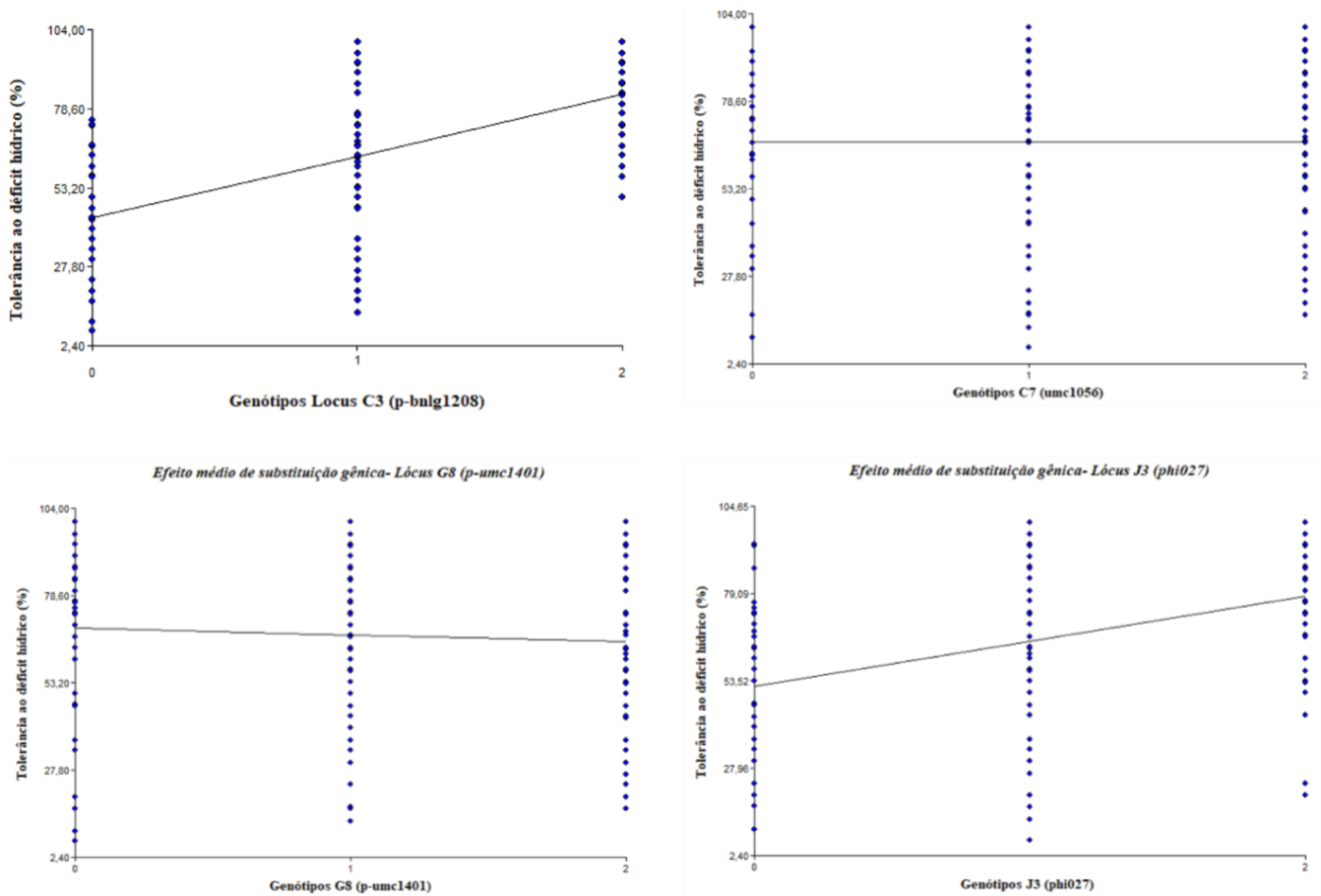
Genótipos com letra diferente, apresentam diferenças estatisticamente significativas ($p < 0.05$).

Fonte: Da autora (2024)

Conforme descrito por Falconer (1996), o coeficiente de regressão (βx) no modelo utilizado, estima o efeito médio da substituição do alelo, ou o efeito médio da substituição de um alelo (suscetível) pelo alelo (tolerante). Para os locos *bnlg1208* e *phi027* encontraram-se valores de βx positivos, de 19.75 e 13.21, respectivamente (Figura 9). Isto indica qual dos alelos parentais aumenta o valor da característica (positivo se P1; negativo se P2).

Para o marcador *bnlg1208* o alelo proveniente do parental 1 (alelo tolerante) incrementa a tolerância das progênies em um 19.75% e no caso do marcador *phi027* o alelo tolerante incrementa a tolerância em um 13.21% nas progênies. No entanto, deve-se levar em conta que até o momento, foi considerada a contribuição de um único loco para o caráter quantitativo da tolerância ao déficit hídrico. Ao considerar vários locos, deve-se também levar em consideração possíveis interações epistáticas entre eles (desvios da epistasia), que juntamente com os desvios da dominância constituem o valor não aditivo do traço (LÓPEZ ANIDO *et al.*, 2020).

Figura 9- Análise de regressão linear. Efeito médio da substituição alélica



Loco	Constante (α)	βx
bnlg1208	43.59	19.75
umc1056	66.85	-0.02
umc1401	66.1	-1.97
Phi027	51.94	13.21

Fonte: Da autora (2024)

A direção da contribuição alélica é quase sempre consistente em estudos onde os resultados para as características medidas, foi encontrada uma correlação altamente positiva, como no presente caso. Adicionalmente, diferentes doadores variam nos efeitos dos alelos em outros locos que influenciam direta ou indiretamente a característica alvo. Isto influencia significativamente o valor dos genótipos parentais que contribuem com os haplótipos componentes mais favoráveis para uma característica determinada. Felizmente, a análise de correlação pode ajudar a determinar qual é o melhor doador (YAN; WARBURTON; CROUCH, 2011).

Por outro lado, os resultados obtidos por meio das análises de contingência, permitiram confirmar a associação de cada loco com o nível de tolerância ao déficit hídrico. Mais uma vez, para os locos bnlg1208 e phi027 houve valores-p estatisticamente significativos ($p < 0.0001$), reforçando os resultados obtidos nas outras análises estatísticas (Tabela 7).

Tabela 7- Frequências absolutas por nível de tolerância ao déficit hídrico.

Loco	Genótipo	Intermediária	Suscetível	Tolerante	Total	gl	Valor-p
bnlg1208	0	14	27	0	41	4	<0.0001
	1	30	17	25	72		
	2	24	0	50	74		
Total		68	44	75	107		
umc1056	0	15	8	18	41	4	0.8421
	1	33	22	30	85		
	2	28	17	33	78		
Total		76	47	81	204		
umc1401	0	18	12	25	55	4	0.9069
	1	31	18	30	79		
	2	27	17	27	71		
Total		76	47	82	205		
	0	23	18	3	44	4	<0.0001
	1	24	23	27	74		
	2	22	3	49	74		
Total		69	44	79	192		

Fonte: Da autora (2024)

Os três tipos de análises estatísticas realizadas, permitiram estabelecer que os locos bnlg1208 e phi027 parecem influenciar a resposta à tolerância ao déficit hídrico das progênies segregantes de milho avaliadas nesta pesquisa.

O loco bnlg1208 mapeado no cromossomo 5 segundo a MaizeGBD, está associada ao gene *fator de transcrição AP2-EREBP-144* e o loco phi027, mapeado no cromossomo 9, está associado ao gene *waxy1* (HARPER *et al.*, 2016).

Devido à informação limitada disponível sobre essas regiões genicas associados às regiões SSR bnlg1208 e phi027 foi realizada uma pesquisa *in silico* para 100 KB *downstream* e

upstream de cada região, com a finalidade de encontrar outros genes associados às respostas de tolerância ao déficit hídrico ou estresses abióticos. Foi possível identificar um total de sete genes candidatos nas proximidades dos dois locos de interesse (Tabela 8). Dois dos sete genes candidatos foram fatores de transcrição. Os genes *fator de transcrição AP2-EREBP-144* e *Heat stress transcription factor C-1 (HsfA1)* associados à região *upstream* do loco *bnlg1208*.

Os fatores de transcrição desempenham um papel crucial na tolerância das plantas ao déficit hídrico, regulando a expressão de genes envolvidos em vários processos fisiológicos, bioquímicos e moleculares que permitem às plantas resistir ao déficit hídrico. Eles ativam ou reprimem a transcrição de vários grupos de genes que codificam proteínas envolvidas em vários aspectos da resposta ao déficit hídrico, como ajuste osmótico, desintoxicação de espécies reativas de oxigênio e síntese de moléculas protetoras como osmoprotetores e antioxidantes, ou codificam para genes aquaporinas, ajudando na regulação da captação, o transporte e a retenção de água em diferentes tecidos vegetais sob condições de déficit hídrico, mantendo assim a homeostase da água, assim como também outros fatores de transcrição atuam como componentes-chave nas vias de sinalização mediadas por hormônios relacionados ao estresse, como o ácido abscísico (ABA) ou atuam na coordenação dos ajustes metabólicos em resposta ao estresse hídrico (CAO *et al.*, 2024; CHENG *et al.*, 2023; HAIDER *et al.*, 2021; LI *et al.*, 2019a; WEERATHUNGA, 2015).

Os fatores de transcrição envolvidos na resposta ao déficit hídrico geralmente interagem com componentes de outras vias de sinalização de estresse, como aquelas ativadas por alta salinidade, temperaturas extremas ou ataque de patógenos. Esta conversa cruzada permite que as plantas integrem múltiplos sinais de estresse e montem respostas apropriadas para lidar com as mudanças nas condições ambientais (RABARA *et al.*, 2014).

Os *AP2/EREBP* (TFs) são uma família de fatores de transcrição essenciais para o crescimento e desenvolvimento adequados das plantas, incluindo a germinação das sementes, o desenvolvimento das raízes e a formação das flores, implicados principalmente na regulação das vias de sinalização hormonal, como o etileno e o ácido abscísico, que são importantes para o desenvolvimento das plantas (LIMA, 2018; LIU *et al.*, 2020b). Esses fatores de transcrição também estão envolvidos na resposta ao estresse abiótico e desempenham um papel central nas respostas de tolerância à seca. Eles têm sido intensamente estudados no contexto da sinalização ABA e da regulação da resposta à seca em várias plantas, porem eles têm sido pouco estudados até agora no milho. Liu *et al.* (2020b), identificaram um novo gene *AP2/EREBP* do milho *ZmERE160* que foi caracterizado e superexpresso em *Arabidopsis*, melhorando

significativamente a tolerância à seca nessa planta modelo (CHENG *et al.*, 2023; LIU *et al.*, 2020b). Também os genes *AP2EREBP* são uns dos principais reguladores do aumento do comprimento da espiga no milho e no controle de desenvolvimento de meristemas, tecidos e órgãos no milho (MENG *et al.*, 2023).

O gene *Heat stress transcription factor C-1 (HsfA1)* pertence à família do fator de choque térmico (*HSF*), o qual parece desempenhar um papel importante na regulação de genes responsivos ao estresse térmico em plantas sob estresse térmico. Além os *HsfA1s* atuam como “reguladores mestres” na rede reguladora do estresse térmico e são necessárias para a ativação de redes transcricionais (HAIDER *et al.*, 2021).

A distribuição da família de genes *HSF* no milho foi analisada por Haider *et al.* (2021) por meio da construção de um mapa cromossômico. Eles concluíram que todos os cromossomas de milho apresentavam pelo menos um gene *HSF*, sugerindo que no ancestral comum mais recente de *Zea mays* os genes *HSF* estão amplamente distribuídos no seu genoma. Por meio da análise filogenética observa-se os *HSFs* de milho são divididos em três classes e ainda em 13 subclasses, o que foi consistente com o número de classes de *HSF* observado em outras monocotiledôneas. Por exemplo, os *HSFs* de *Oryza sativa*, *Brachypodium distachyon*, *Sorghum* e *Triticum aestivum* também são divididos em três classes e 13 subclasses. Apesar disso, existem diferenças entre os números de *HSF* presentes nas diferentes monocotiledôneas (HAIDER *et al.*, 2021).

Li *et al.* (2019c) isolaram o *ZmHsf05* do milho utilizando métodos de clonagem homóloga. O *ZmHsf05* foi expresso em muitos tecidos de milho e seu nível de expressão foi aumentado pelo tratamento de estresse térmico. Também foi testado em mudas de *Arabidopsis* com superexpressão de *ZmHsf05* que aumentaram tanto a termotolerância basal quanto a adquirida. Após o estresse térmico, as linhagens superexpressoras de *ZmHsf05* apresentaram maior taxa de sobrevivência e teor de clorofila em comparação com as plântulas silvestres (controle). Estes resultados sugerem que o *ZmHsf05* desempenha um papel importante na termotolerância basal e adquirida nas plantas e pode ser testado para a tolerância a outros tipos de estresses abióticos como o déficit hídrico.

Por outro lado, no sentido *downstream* do loco bnlg1208 foi encontrado o gene candidato *3-ubiquitin-protein ligase CCNB1IP1* (Tabela 8; Figura 10). O gene *E3 ubiquitin-protein ligase* desempenha um papel crucial no crescimento e desenvolvimento das plantas, ajudando na síntese de flavonóides e lignina, que são importantes para as respostas das plantas ao estresse abiótico. Em tabaco transgênico, a superexpressão do gene *Fm4CL-like 1*, uma *E3 ligase*,

aumentou o conteúdo de lignina e engrossou a parede celular do xilema da planta, além de apresentar menor teor de H₂O₂ e maiores atividades de peroxidase (POD) e superóxido dismutase (SOD), indicando um papel na regulação de ROS e levando a uma melhor resistência à seca e à tolerância ao estresse osmótico em comparação com o tipo selvagem (LI *et al.*, 2019b). As *E3 ligases* são importantes para manter a homeostase proteica, regulando processos celulares como crescimento, diferenciação, apoptose, secreção, endocitose, transcrição gênica, expressão, transdução de sinal e resposta imune por isso descobriu-se que as *E3 ligases* desempenham papéis essenciais na modulação da resposta a estressores bióticos, abióticos e hormonais (LI *et al.*, 2022; WANG *et al.*, 2022).

Moin *et al.* (2021) estudaram os níveis de expressão de um grande conjunto de *E3 ligases* de ubiquitina e cumarato-Co-A (*CCLs*) em plantas de arroz. As proteínas ligases, como *HUB*, *CCNB1IP*, *ZFP* e ligas do tipo *EL RING* foram altamente expressas e mostraram estar envolvidas na regulação do ciclo celular e no cruzamento meiótico, que desempenham papéis importantes na organogênese e gametogênese e, portanto, na folha, no caule, morfogênese radicular e floral.

Li *et al.* (2022) identificaram um total de 79 genes *E3 ligases* no genoma do milho, que foram classificados em sete categorias semelhantes a outras plantas como arroz, soja, cevada, citrinos e banana. Pela distribuição cromossômica desses genes foi verificado que eles estavam distribuídos de forma desigual dentro de cada cromossomo, com o maior número de genes encontrados nos cromossomos 1, 4 e 5. Os perfis de expressão da família de genes *E3 ligases* do milho apresentaram características de especificidade tecidual e resposta ao estresse abiótico e ao tratamento hormonal.

O loco phi027, localizado no cromossomo 9, está associado ao gene *Waxy1* e a mais quatro genes que flanqueiam essa região.

O gene *Waxy1*, com 3.93 kb de comprimento e contendo 14 exons no braço longo do cromossomo 9, codifica o *amido sintase ligado a grânulos (GBSS)*, que controla a síntese de amilose no endosperma do milho (ZHANG *et al.*, 2013). Embora o gene *Waxy1* (tipo selvagem dominante) possa converter ADP-glicose em amilose, o gene *waxy1* (mutante recessivo) inibe enormemente a conversão em amilose, o que leva ao aumento do acúmulo de amilopectina. Alelos homocigotos recessivos *waxy* resultam no endosperma do milho ceroso com uma textura pegajosa e compreendendo quase apenas amilopectina (99%) (SA *et al.*, 2023). Esta

composição alterada do amido resulta em grãos que têm uma aparência de cera de vela, portanto, é conhecido como “milho ceroso”. O amido de milho feito de milho ceroso tem diversas aplicações. É usado para melhorar a uniformidade, estabilidade e fornecer textura a diversos produtos alimentícios. O amido de milho ceroso também tem aplicações nas indústrias têxteis e de adesivos e é utilizado como aglutinante em processos de fabricação na indústria de papel (CHILCOAT; LIU; SANDER, 2017). Embora o papel dos genes *Waxy* não tenha sido associado à tolerância ao déficit hídrico, sabe-se que o estresse hídrico influencia nos níveis de expressão dos genes *Waxy* como foi reportado por Chen *et al.* (2017), que demonstraram que a redução na expressão gênica ou nível de transcrição, expressão proteica e níveis de fosforilação de enzimas relacionadas à biossíntese de amido sob condições de déficit hídrico foram responsáveis pela diminuição significativa no teor total de amido e no rendimento de grãos.

Dos outros quatro genes candidatos associados ao loco *phi027*, o gene *nuclear speckle-RNA binding protein A* foi localizado em sentido *downstream* do loco. As proteínas de ligação ao RNA *speckle nuclear* (*NSRs*) são uma família de *RBP*s (proteínas de ligação ao RNA) que atuam como reguladores do *Splicing* Alternativo (*AS*) e processos de desenvolvimento regulados por auxina, como a formação de raízes laterais em *Arabidopsis thaliana* (LUCERO *et al.*, 2020). Os *RBP*s identificados funcionam como acompanhantes de RNA que governam muitos aspectos do metabolismo do RNA, incluindo processamento/edição de pré-mRNA, transporte, estabilidade/decaimento e tradução. Tendo um papel no *splicing* alternativo, os *RBP*s provavelmente modularão mais de 60% do transcriptoma da planta que regula o crescimento, o desenvolvimento e as respostas ao estresse das plantas, o que indica a importância dos *RBP*s. Portanto, os *RBP*s são cada vez mais reconhecidos como reguladores pós-transcricionais e estão sendo utilizados para melhorar características agronômicas importantes, incluindo rendimento, biomassa e tolerância ao estresse (MUTHUSAMY *et al.*, 2021). No entanto, as funções dos *RBP*s nas plantas permanecem pouco compreendidas (ZUO *et al.*, 2019).

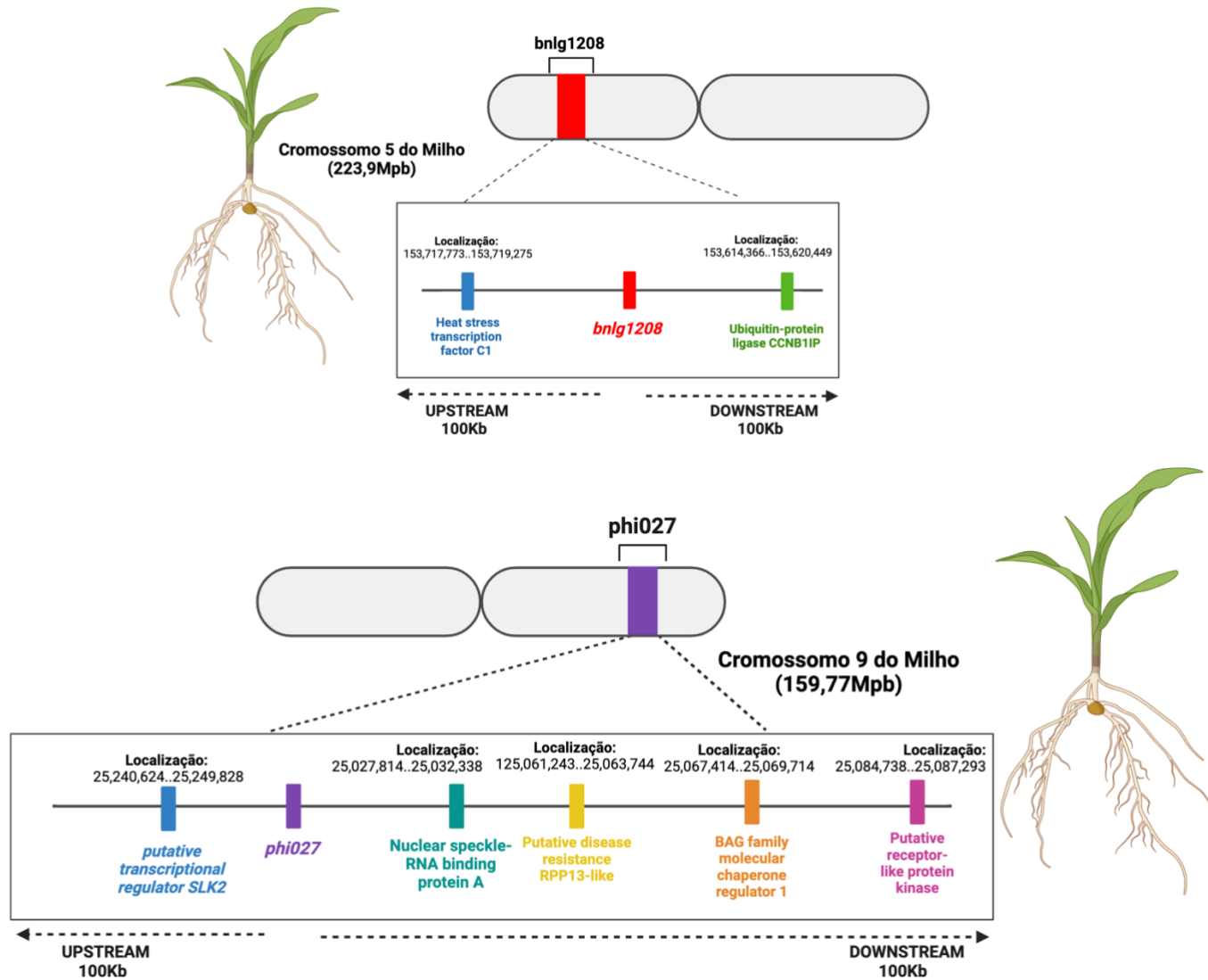
Zuo *et al.* (2019) clonaram o gene *dek42* (*RBP*) localizado no núcleo e expresso constitutivamente em vários tecidos de milho. Eles clonaram o *Dek42* por isolamento de *tag Mutator* e posteriormente confirmado pela técnica de repetições palindrômicas curtas regularmente interespaçadas (CRISPR). A mutação *Dek42* causou uma redução significativa no acúmulo da proteína DEK42 nos grãos mutantes e a análise de RNA-seq mostrou que foi alterada a expressão de milhares de genes durante o desenvolvimento do grão de milho e o *splicing* alternativo em genes expressos.

Tabela 8- Genes candidatos identificados nas proximidades 100KB dos locos bnlg1208 e phi027.

Gene	Localização	Tamanho	mRNA	BLAST nr	Posição relativa	SSR	Cromossomo
ZEAMMB73_Zm00001d016254	153,614,366..153,620,449	6.084	<i>mRNA-E3 ubiquitin-protein ligase CCNB1IP1-like protein</i>	CM000781.4 (153,614,366..153,620,449)	<i>downstream</i>	bnlg1208	5
ZEAMMB73_Zm00001d016255	153,717,773..153,719,275	1.503	<i>mRNA-Heat stress transcription factor C-1</i>	CM000781.4 (153,717,773..153,719,275)	<i>Upstream</i>	bnlg1208	5
ZEAMMB73_Zm00001d045509	25,027,814..25,032,338	4.525	<i>mRNA nuclear speckle- RNA binding protein A</i>	CM000785.4 (25,027,814..25,032,338)	<i>Downstream</i>	phi027	9
ZEAMMB73_Zm00001d045512	25,061,243..25,063,744	2.502	<i>mRNA-Putative disease resistance RPP13-like protein 3</i>	CM000785.4 (25,061,243..25,063,744)	<i>Downstream</i>	phi027	9
ZEAMMB73_Zm00001d045513	25,067,414..25,069,714	2.301	<i>mRNA-BAG family molecular chaperone regulator 1</i>	CM000785.4 (25,067,414..25,069,714)	<i>Downstream</i>	phi027	9
ZEAMMB73_Zm00001d045515	25,084,738..25,087,293	2.556	<i>mRNA-Putative receptor-like protein kinase family protein</i>	CM000785.4 (25,084,738..25,087,293)	<i>Downstream</i>	phi027	9
ZEAMMB73_Zm00001d045516	25,240,624..25,249,828	9.205	<i>mRNA-putative transcriptional regulator SLK2</i>	CM000785.4 (25,240,624..25,249,828)	<i>Upstream</i>	phi027	9

Fonte: Da autora (2024)

Figura 10- Representação gráfica dos cromossomos 5 e 9 mostrando os genes encontrados nas proximidades dos locos p-bnlg1208 e phi027.



Fonte: Da autora (2024)

Outro gene candidato localizado nas proximidades do loco phi027 no sentido *downstream* é o *putative disease resistance RPP13-like protein 3 (ZmRPP13-LK3)*. O *ZmRPP13-LK3* é uma nova adenilil ciclase (AC) em milho que catalisa a conversão de *ATP* para *cAMP*. Está envolvido na resistência ao calor regulada por ácido abscísico (ABA). Em condições de estresse térmico, as concentrações de *ZmRPP13-LK3* e *cAMP* no mutante vp5 deficiente em ABA foram significativamente menores do que aquelas no tipo selvagem. Além, foi descoberto que a aplicação de *8-Br-cAMP*, um análogo de *cAMP*, aumentou a expressão induzida pelo calor de proteínas de choque térmico em plantas selvagens e aliviou o estresse oxidativo ativado pelo calor. Estas descobertas sugerem que *ZmRPP13-LK3* e *cAMP* desempenham um papel na resistência mediada pelo ácido abscísico (ABA) ao estresse térmico no milho indicando que o *ZmRPP13-LK3*, contribui para a resposta mediada pelo ABA ao estresse térmico no milho, aumentando potencialmente a resistência ao calor (YANG *et al.*, 2021). Em contraste, em um estudo feito por Korsman (2015) no qual foram avaliadas as respostas moleculares do milho ao seu patógeno foliar, *Cercospora zeina*, concluíram que genes como o *RPP13* entre outros, podem estar envolvidos na resposta de resistência ao patógeno exibindo maior expressão constitutiva nas plantas resistentes do que nas plantas suscetíveis.

Também foi encontrado o gene candidato *chaperone regulator 1* nas proximidades do loco phi027 no sentido *downstream*. Ele pertence à família molecular *BAG* (atogene associado a Bcl-2). A família *BAG* são reguladores acompanhantes que modular uma série de processos diversos que vão desde a proliferação à parada do crescimento e morte celular. Essas proteínas se distinguem por um domínio *BAG* conservado que interage diretamente com *Hsp70* e proteínas *Hsc70* para regular sua atividade (DOUKHANINA *et al.*, 2006).

Os homólogos de *BAG* em plantas não foram caracterizados na mesma extensão que em mamíferos e relativamente pouco se sabe sobre a função de proteínas semelhantes a *BAG* em plantas. No entanto, o uso de novas ferramentas bioinformáticas em busca de sequências conservadas do domínio *BAG* das proteínas *BAG* humanas, permitiram identificar oito genes que codificam para domínios conservados *BAG* no genoma de *A. thaliana* com alta significância estatística. Neste estudo realizado por Yan, He e Zhang (2003), estes genes foram localizados nos cromossomos 1, 2, 3 e 5 e foi possível agrupar esses oito genes em três classes distintas, consistindo principalmente em diferenças nos motivos dos domínios.

Os genes do tipo *BAG* parecem estar amplamente distribuídos nas plantas. De fato, vários marcadores de sequência expressos (*ESTs*) de supostos genes *BAG* de diferentes plantas expostas a estresses bióticos e abióticos, apresentaram *ESTs* associados à expressão específica em tecido. Esta observação sugere que os genes *BAG* das plantas estão envolvidos nas respostas ambientais e de desenvolvimento, pois regulam processos semelhantes à apoptose que vão do ataque de patógenos ao estresse abiótico (DOUKHANINA *et al.*, 2006; YAN; HE; ZHANG, 2003).

Também foi encontrado o gene candidato *Putative receptor-like protein kinase* próximo ao loco phi027 no sentido *downstream*. As proteínas quinases estão envolvidas na sinalização de estresse em sistemas vegetais e animais. O hormônio ácido abscísico está associado às respostas das plantas a estresses como seca, salinidade e frio (EDEN *et al.*, 2002). Em diversos estudos tem sido observado que proteínas quinases ativadas por ácido abscísico (*AAPK*), controlam a abertura de estômatos e regulam os canais iônicos da membrana plasmática (KIM *et al.*, 2010). No entanto, as plantas têm muitos outros tipos de quinases envolvidos em estresses, cerca de 4% de todos os genes de *Arabidopsis thaliana* codificam proteínas quinases típicas, e o genoma do arroz codifica 40% a mais (CHEN *et al.*, 2021).

Basicamente para todos os tipos de proteínas quinases já descobertas, têm sido associadas ao papel importante na tolerância ao déficit hídrico ou outros estresses abióticos. Tanto as proteínas *MAPK* (*Mitogen-activated protein kinase*), as *CPKs* (*Calcium-dependent protein kinases*), e as *RLK* (*Receptor-like protein kinases*) estão envolvidos na sinalização ABA e no déficit hídrico (CHEN *et al.*, 2021).

Um estudo de expressão de um gene da família dos *RLK* (*Receptor-like protein kinases*), permitiu concluir que níveis altos de expressão do gene *SbER2-1* em folhas e caules sob déficit hídrico, aumenta a tolerância à seca. Tanto em *Arabidopsis* como no milho, a expressão desse gene aumenta a tolerância à seca, particularmente em termos de eficiência no uso da água e taxa fotossintética líquida. Neste estudo também foi concluído que o *SbER2-1* está envolvido na regulação dos sistemas fotossintéticos e no metabolismo dos fenilpropanóides, que são mecanismos importantes para a tolerância ao déficit hídrico (LI *et al.*, 2019a).

Na busca de outros genes próximos à região do loco phi027, foi encontrado um gene no sentido *upstream*, sugerido como *putative transcriptional regulator SLK2*. Os reguladores *SLK* são genes responsivos ao estresse que são controlados negativamente sob condições normais

de crescimento que são o principal mecanismo utilizado pelas plantas no silenciamento de genes (WEERATHUNGA, 2015).

Os genes sob repressão e ativação transcricional são controlados epigeneticamente, em que o nível de transcrição depende da acessibilidade da sequência codificadora da proteína à maquinaria de transcrição, e não da informação no próprio DNA. A metilação do DNA e a organização do DNA em uma estrutura de heterocromatina altamente compacta são marcas de silenciamento de genes epigenéticos. Para genes sob silenciamento epigenético, a estrutura da cromatina é controlada dinamicamente por diferentes tipos de complexos proteicos (LAW; JACOBSEN, 2010).

Plantas mutantes no sistema LUH (LEUNIG- HOMOLOG) e os genes *SLK1* e *SLK2* apresentam elevada tolerância ao estresse salino e osmótico e níveis mais elevados de expressão do gene responsivo ao estresse abiótico em condições sem estresse. Além disso, o LUH interage fisicamente com as histonas H2B e H3 e regula direta ou indiretamente a estrutura da cromatina nos genes responsivos ao estresse abiótico. Esses dados fornecem uma visão sobre os novos papéis de LUH, *SLK1* e *SLK2* na regulação gênica da resposta ao estresse abiótico e iluminam a função de LUH na remodelação da cromatina (WEERATHUNGA, 2015).

Os resultados obtidos pela análise *in silico*, demonstram de que ainda existe pouca informação sobre os genes e fatores da maquinaria molecular envolvidos nas respostas de tolerância ao déficit hídrico e que esta área de estudo está em constante crescimento. Isso sugere que os sete genes encontrados neste estudo, podem ser alvo para futuras pesquisas sobre tolerância ao déficit hídrico, e também possíveis marcadores nos processos de seleção em programas de melhoramento de milho usando o desenvolvimento de cultivares tolerantes ao déficit hídrico.

O estudo dos mecanismos de regulação genética é uma área de pesquisa ativa, e novas descobertas podem refinar continuamente nossa compreensão de regiões genômicas específicas, que desempenham papéis importantes nas respostas de tolerância ao déficit hídrico no milho.

4. CONCLUSÕES

A segregação da população F₂, se ajustou ao modelo de segregação mendeliana (1:2:1) e permitiu estabelecer que os locos bnlg1208 e phi027 estão associados à resposta a tolerância ao déficit hídrico e são potenciais marcadores moleculares para seleção precoce assistida por marcadores (MAS) em programas de melhoramento de milho.

A presença dos alelos provenientes dos parentais tolerantes nas populações estudadas para os locos bnlg1208 e phi027, influenciam no aumento da tolerância.

Dos sete genes associados aos locos bnlg1208 e phi027, seis pertencem aos grupos de genes característicos de respostas de tolerância ao déficit hídrico e podem ser novos alvos de estudo para a seleção assistida.

REFERÊNCIAS

- ADEBAYO, M. A. *et al.* Diversity assessment of drought tolerant exotic and adapted maize (*Zea mays* L.) inbred lines with microsatellite markers. **Journal of Crop Science and Biotechnology**, v. 18, n. 3, p. 147–154, 1 set. 2015.
- AUOBI, A.; RAHMANI, F. The Effects of Drought and Salt Stresses on the Antioxidant System and PIPs Gene Expression in *Zea mays* L. **Jordan Journal of Agricultural Sciences**, v. 13, n. 4, p. 997–1013, 9 nov. 2017.
- BALZARINI, M. *et al.* **Análisis de Datos de Marcadores con Info-Gen**. [s.l.] Universidad Nacional de Córdoba-CONICET, 2006.
- BALZARINI, M.; DI RIENZO, J. **Info-Gen: Software para análisis estadístico de datos genéticos**. Facultad de Ciencia Agropecuarias. Universidad Nacional de Córdoba. Argentina, 1 jan. 2003.
- BELALIA, N. *et al.* Analysis of genetic diversity and population structure in Saharan maize (*Zea mays* L.) populations using phenotypic traits and SSR markers. **Genetic Resources and Crop Evolution**, v. 66, n. 1, p. 243–257, 31 jan. 2019.
- BOCIANOWSKI, J. *et al.* Identification of Associations between SSR Markers and Quantitative Traits of Maize (*Zea mays* L.). **Agronomy**, v. 11, n. 1: 182, 2021.
- BROMAN, K. W.; SAUNAK, S. **A Guide to QTL Mapping with R/qtl**. Springer ed. New York: [s.n.]. v. Vol. 46, 2009.
- CAO, L. *et al.* Transcription factor ZmDof22 enhances drought tolerance by regulating stomatal movement and antioxidant enzyme activities in maize (*Zea mays* L.). 2024.
- CHEN, GX. *et al.* In vivo phosphoproteome characterization reveals key starch granule-binding phosphoproteins involved in wheat water-deficit response. **BMC Plant Biol**, v. 17, p. 1-13, 2017.
- CHEN, X. *et al.* Protein kinases in plant responses to drought, salt, and cold stress. **Journal of Integrative Plant Biology**, Blackwell Publishing Ltd, 1 jan. 2021.
- CHENG, C. *et al.* Wide-Range Portrayal of AP2/ERF Transcription Factor Family in Maize (*Zea mays* L.) Development and Stress Responses. **Genes**, v. 14, n. 1, 1 jan. 2023.
- CHILCOAT, D.; LIU, Z. BIN; SANDER, J. Use of CRISPR/Cas9 for Crop Improvement in Maize and Soybean. Em: **Progress in Molecular Biology and Translational Science**, [s.l.] Elsevier B.V, v. 149, p. 27–46, 2017.
- CHRISTY SANTOS, M. **Expressão de genes relacionados à tolerância ao estresse hídrico em sementes e em tecidos de plântulas de milho**. Lavras, MG: Universidade Federal de Lavras, 2016.

DOUKHANINA, E. V. *et al.* Identification and functional characterization of the BAG protein family in *Arabidopsis thaliana*. **Journal of Biological Chemistry**, v. 281, n. 27, p. 18793–18801, 7 jul. 2006.

EDEN, S. *et al.* Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck. **Nature**, v. 418, n. 6899, p. 790–793, 15 ago. 2002.

EMBRAPA CLIMA TEMPERADO. **Documentos 218: Bulks de DNA na Bulks de DNA na caracterização de caracterização de germoplasma vegetal**. Pelotas, RS: [s.n.]. Disponível em: <www.cpact.embrapa.br>. Acesso em: 9 dez. 2023.

FALCONER, D. S. **Introduction to quantitative genetics**. [s.l.] Pearson Education India, 1996.

FAO. **OECD-FAO Agricultural Outlook 2023-2032**. [s.l.: s.n.]. Disponível em: <<http://dx.doi.org/10.1787/agr-outl-data->>. Acesso em: 10 dez. 2023.

FAROOQI, M. Q. U. *et al.* Bulk segregant analysis (BSA) for improving cold stress resistance in maize using SSR markers. **Genetics and Molecular Research**, v. 15, n. 4, 19 dez. 2016.

FORBES AGRO. **O estouro do milho brasileiro: a supersafra nunca foi tão grande**. [s.l.: s.n.]. Disponível em: <<https://forbes.com.br/forbesagro/2023/06/o-estouro-do-milho-brasileiro-a-supersafra-nunca-foi-tao-grande/>>. Acesso em: 15 jan. 2024.

FRANCIS, D. M.; HEATHER, L. M.; ROBBINS, M. **Analysis of Variance for Plant Breeding**. Disponível em: <https://plant-breeding-genomics.extension.org/analysis-of-variance-for-plant-breeding/>. Acesso em: 14 dez. 2023.

GARRIDO-CARDENAS, J. A.; MESA-VALLE, C.; MANZANO-AGUGLIARO, F. Trends in plant research using molecular markers. **Planta**, Springer Verlag, v. 247, n. 3, p. 543-557, 2018.

HAIDER, S. *et al.* In silico characterization and expression profiles of heat shock transcription factors (Hsfs) in maize (*zea mays* l.). **Agronomy**, v. 11, n. 11, 1 nov. 2021.

HARPER, L. *et al.* Maizegdb: The maize genetics and genomics database. **Methods in Molecular Biology**, v. 1374, p. 187–202, 2016.

HASAN, S. A. *et al.* Water use efficiency in the drought-stressed sorghum and maize in relation to expression of aquaporin genes. **Biologia Plantarum**, v. 61, n. 1, p. 127–137, 2017.

HE, C. *et al.* Early drought-responsive genes are variable and relevant to drought tolerance. **G3: Genes, Genomes, Genetics**, v. 10, n. 4, p. 1657–1670, 1 abr. 2020.

HU, H.; XIONG, L. Genetic Engineering and Breeding of Drought-Resistant Crop. **Annual Review of Plant Biology**, v. 65, p. 715–741, 2014.

KALINOWSKI, S. T.; TAPER, M. L.; MARSHALL, T. C. Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. **Molecular Ecology**, v. 16, n. 5, p. 1099–1106, 1 mar. 2007.

KIM, T. H. *et al.* Guard cell signal transduction network: Advances in understanding abscisic acid, CO₂, and Ca²⁺ signaling. **Annual Review of Plant Biology**, v. 61, p. 561–591, 2 jun. 2010.

KIZIS, D.; PAGÈS, M. Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway. **Plant Journal**, v. 30, n. 6, p. 679–689, 2002.

KLEIN, H. S.; LUNA, F. V. The Impact of the Rise of Modern Maize Production in Brazil and Argentina. **Historia Agraria**, n. 86, p. 1–38, 1 abr. 2022.

KORSMAN, J. N. **Molecular responses of maize to its foliar pathogen, Cercospora zeina**. University of Pretoria, 2015.

KUMAR, R. *et al.* Diversity analysis of maize inbred lines on the basis of morphological and simple sequence repeat markers. **Ecology, Environment and Conservation Journal**, v. 22, p. S147–S153, jun. 2016.

KUMARI, A. *et al.* Genetic diversity analysis in maize (*Zea mays* L.) using SSR markers. **Journal of Pharmacognosy and Phytochemistry**, p. 1, 2018.

LAW, J. A.; JACOBSEN, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. **Nature Reviews Genetics**, v. 11, p. 204–220, 2010.

Li, G. L. *et al.* ZmHsf05, a new heat shock transcription factor from *Zea mays* L. improves thermotolerance in *Arabidopsis thaliana* and rescues thermotolerance defects of the *athsf2* mutant. **Plant science**, v. 283, p. 375–384, jun. 2019.

LI, H. *et al.* A leucine-rich repeat-receptor-like kinase gene SbER2-1 from sorghum (*Sorghum bicolor* L.) confers drought tolerance in maize. **BMC Genomics**, v. 20, n. 1, 15 out. 2019a.

LI, X. *et al.* Classification and Expression Profile of the U-Box E3 Ubiquitin Ligase Enzyme Gene Family in Maize (*Zea mays* L.). **Plants**, v. 11, n. 19, 1 out. 2022.

LI, Y. *et al.* Genome-wide identification, phylogenetic and expression analysis of the maize HECT E3 ubiquitin ligase genes. **Genetica**, v. 147, n. 5–6, p. 391–400, 1 dez. 2019b.

LIANG, Z. *et al.* Effects of electrophoresis buffer, electrophoresis time, and annealing temperature on the quality of SSR molecular detection in rice varieties. **Cogent Food and Agriculture**, v. 9, n. 2, 2023.

LIMA MARQUES, T. **Expressão gênica em sementes, espigas e plântulas de linhagens de milho associada à tolerância ao déficit hídrico**. Lavras, MG: Universidade Federal de Lavras, 2018.

- LIU, S. *et al.* Mapping regulatory variants controlling gene expression in drought response and tolerance in maize. **Genome Biology**, v. 21, n. 1, 6 jul. 2020a.
- LIU, W. *et al.* The Maize AP2/EREBP Transcription Factor ZmEREB160 Enhances Drought Tolerance in Arabidopsis. **Tropical Plant Biology**, v. 13, n. 3, p. 251–261, 1 set. 2020b.
- LONG, T. *et al.* Genome-wide identification of ZmSnRK2 genes and functional analysis of ZmSnRK2.10 in ABA signaling pathway in maize (*Zea mays* L). **BMC Plant Biology**, v. 21, n. 1, 1 dez. 2021.
- LÓPEZ ANIDO, F. *et al.* **Unidad 4: Genética Cuantitativa**. Rosario, Argentina: [s.n.].
- LUCERO, L. *et al.* Evolution of the small family of alternative splicing modulators nuclear speckle RNA-binding proteins in plants. **Genes**, v. 11, n. 2, 1 fev. 2020.
- MARSHALL, T. C. *et al.* Statistical confidence for likelihood-based paternity inference in natural populations. **Molecular Ecology**, v. 7, n. 5, p. 639–655, 1 maio 1998.
- MASON, A. S. Ssr genotyping. **Methods in Molecular Biology**, v. 1245, p. 77–89, 2015.
- MENG, S. *et al.* Transcriptomic analysis reveals the regulation of early ear-length development in maize. **Plant Growth Regulation**, v. 100, n. 1, p. 97–105, 1 maio 2023.
- MIKIĆ, S. *et al.* Agronomic and molecular evaluation of maize inbred lines for drought tolerance. **Spanish Journal of Agricultural Research**, v. 14, n. 4, p. 1-13, 2016.
- MITCHELSON, K. R.; CHENG, J. **Capillary electrophoresis of nucleic acids**. Springer Science ed. [s.l.: s.n.]. v. Vol. 162
- MOIN, M. *et al.* Constitutive expression of Ribosomal Protein L6 modulates salt tolerance in rice transgenic plants. **Gene**, v. 789, 15 jul. 2021.
- MORETTI, S. In Silico Experiments in Scientific Papers on Molecular Biology. **Science & Technology Studies**, v. 24, n. 2, p. 23–42. 2011.
- MUTHUSAMY, M. *et al.* Plant rna binding proteins as critical modulators in drought, high salinity, heat, and cold stress responses: An updated overview. **International Journal of Molecular Sciences**, v.22, n.13, 1 jul. 2021.
- NAEEM, R. Molecular Markers in Plant Genotyping. **Journal of Bio-Molecular Sciences (JBMS)**, v.2, n. 3, p. 78-85, 30 oct. 2014.
- NAEEM, R.; MIRZA, B. High-Throughput DNA Extraction and Optimization of PCR Efficiency for Barley SSRs Genotyping. **Arabian Journal for Science and Engineering**, v. 43, n. 1, p. 143–154, 1 jan. 2018.
- NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics**, v. 89, n. 3, p. 583-90, jul. 1978.
- PARISSEAUX, B.; BERNARDO, R. In silico mapping of quantitative trait locos in maize. **Theoretical and Applied Genetics**, v. 109, n. 3, p. 508–514, 2004.

- PEER, L. A. *et al.* Genetic Diversity Analysis in Maize Landraces under Temperate Ecology. **Agricultural Science Digest**, v. 42, n. 5, p. 541–547, 1 out. 2022.
- QUARRIE STEVE A *et al.* Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. **Journal of Experimental Botany**, v. 50, n. 337, p. 1299–1306, 1999.
- QUADRAT ULLAH FAROOQI, M. *et al.* Bulk Segregant Analysis (BSA) for the improvement of Drought Resistance in Maize (*Zea mays* L.) Inbred Lines as revealed by SSR Molecular Markers. **Article in Research Journal of Biotechnology**, v. 13, n. 2, 2018.
- R DEVELOPMENT CORE TEAM (2019). **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2019. Disponível em: <http://www.R-project.org/>. Acesso em: 21 fev. 2021.
- RABARA ROEL C.; TRIPATHI PRATEEK.; RUSHTON PAUL J. The Potential of Transcription Factor-Based Genetic Engineering in Improving Crop Tolerance to Drought. **OMICS: A Journal of Integrative Biology**, v. 18, n. 10, p. 601-614, 2014.
- RIGHETTI, P. G.; GELFI, C.; D'ACUNTO, M. R. Recent progress in DNA analysis by capillary electrophoresis. **ELECTROPHORESIS**, v. 23, n. 10, p. 1361–1374, 1 maio 2002.
- ROMERO NAVARRO, J. A. *et al.* A study of allelic diversity underlying flowering-time adaptation in maize landraces. **Nature Genetics**, v. 49, n. 3, p. 476–480, 1 mar. 2017.
- SA, K. J. *et al.* Association Mapping of Amylose Content in Maize RIL Population Using SSR and SNP Markers. **Plants**, v. 12, n. 2, 1 jan. 2023.
- SALGADO, K. *et al.* Mapping of quantitative trait locus associated with maize tolerance to high seed drying temperature. **Genetics and Molecular Research**, v. 7, n. 4, p. 1319–1326, 2008.
- SALGADO, K. C. **Mapeamento de QTLs associados à tolerância à alta temperatura de secagem em sementes de milho**. Lavras, MG: Universidade Federal de Lavras, 2005.
- SCHOBBER, P.; SCHWARTE, L. A. Correlation coefficients: Appropriate use and interpretation. **Anesthesia and Analgesia**, v. 126, n. 5, p. 1763–1768, 1 maio 2018.
- SHARMA, T. *et al.* Molecular characterization and genetic diversity analysis of selected maize inbreds using SSR markers. **Journal of Environmental Biology**, v. 39, n. 2, p. 228–236, 1 mar. 2018.
- SHIMELIS, H. *et al.* Genetic diversity analysis of elite maize inbred lines of diverse sources using SSR markers. **Maydica**, v. 60, 2015.
- SHIRI, M. Identification of informative simple sequence repeat (SSR) markers for drought tolerance in maize. **African Journal of Biotechnology**, v. 10, n. 73, p. 16414–16420, 21 nov. 2011.

SLAVOV, G. T. *et al.* Highly variable SSR markers in Douglas-fir: Mendelian inheritance and map locations. **Theoretical and Applied Genetics**, v. 108, n. 5, p. 873–880, mar. 2004.

SOARES, N. M. **Análise in silico de genes que participam das respostas aos estresses abióticos (seca/salinidade) nos genomas de plantas superiores**. Recife, PE: Universidade Federal de Pernambuco, 2012.

TALAAT, N. B.; IBRAHIM, A. S.; SHAWKY, B. T. Enhancement of the Expression of ZmBZR1 and ZmBES1 Regulatory Genes and Antioxidant Defense Genes Triggers Water Stress Mitigation in Maize (*Zea mays* L.) Plants Treated with 24-Epibrassinolide in Combination with Spermine. **Agronomy**, v. 12, n. 10, 1 out. 2022.

TAVARES, M. F. M. Mecanismos de separação em eletroforese capilar. **Química nova**, v. 20, n. 5, p. 493-51. 1997.

USDA. **World Agricultural Production**. Disponível em: <https://fas.usda.gov/data/world-agricultural-production>. Acesso em: 2023.

VILLANUEVA-MEJÍA, D. *et al.* Genetic differentiation of *Tecia solanivora* from Colombia based on two mitochondrial genes: Cytochrome b and cytochrome oxydase I. **Journal of Insect Science**, 2015.

VON PINHO, I. **Fenotipagem e expressão gênica para tolerância ao déficit hídrico em milho**. Tese—Lavras, MG: Universidade Federal de Lavras (UFLA), 2021.

WANG, S.; LV, X.; ZHANG, J.; CHEN, D.; CHEN, S.; FAN, G.; MA, C.; WANG, Y. Roles of E3 Ubiquitin Ligases in Plant Responses to Abiotic Stresses. **International Journal of Molecular Sciences**, v. 23. 2022

WEERATHUNGA ARACHCHILAGE ACHIRA S. **A Novel Transcription Factor in Arabidopsis thaliana Abiotic Stress Response**. Thesis—University of New Orleans, 2015.

YAN, J.; HE, C.; ZHANG, H. The BAG-family proteins in *Arabidopsis thaliana*. **Plant Science**, v. 165, n. 1, p. 1-7, jul. 2003.

YAN, J.; WARBURTON, M.; CROUCH, J. Association Mapping for Enhancing Maize (*Zea mays* L.) Genetic Improvement. **Crop Science**, v.51, mar-abr. 2011

YANG, H. *et al.* A new adenylyl cyclase, putative disease-resistance RPP13-like protein 3, participates in abscisic acid-mediated resistance to heat stress in maize. **Journal of Experimental Botany**, v. 72, n. 2, p. 283–301, 2 fev. 2021.

ZHANG, W. *et al.* Increasing Lysine Content of Waxy Maize through Introgression of Opaque-2 and Opaque-16 Genes Using Molecular Assisted and Biochemical Development. **PLoS ONE**, v. 8, n. 2, 15 fev. 2013.

ZHANG, Y. *et al.* Identification of the Maize LEA Gene Family and Its Relationship with Kernel Dehydration. **Plants**, v. 12, n. 21, 1 nov. 2023.

ZHONG, M.-S. *et al.* MdCER2 conferred to wax accumulation and increased drought tolerance in plants. **Plant Physiology and Biochemistry**, v. 149, p. 277–285, 2020.

ZUO, Y.; FENG, F.; QI, W.; SONG, R. Dek42 encodes an RNA-binding protein that affects alternative pre-mRNA splicing and maize kernel development. **Journal of Integrative Plant Biology**, v. 61, n. 6, p. 728-748, jun. 2019.

MATERIAL ANEXO

Tabela S1 - Marcadores SSR selecionados no MaizeGBD

Numero	Nome do Loco	Primer F	Primer R	Tipo de Repetição	Cromossomo
1	bnlg1083	ACAGTCTGTTGGGAACAGG	CAACGCTGGTTTGTGCTTTA	AG(29)	1
2	bnlg1297	TCTCGATCGCTCCGATCTAT	GACTCAACTCCAAAAGGCGA	AG(32)	2
3	bnlg1614	CCAACCCACCCAGAGGAGA	AGCGGGCGAGATCTTCAT	AG(15)	1
4	bnlg1866	CCCAGCGCATGTCAACTCT	CCCCGGTAATTCAGTGGATA	AG(11)	1
5	phi053	CTGCCTCTCAGATTCAGAGATTGAC	AACCCAACGTAATCCGGCAG	ATAC	3
6	umc1056	CGGATCGCTTTTACCGTCTATAA	AGCAAGAGTAGCGTTCATTTCAG	(AGCA)4	5
7	bnlg2190	TCCTCCTTCATCCCCTTCTT	CCCAGTATCATTGCCCAATC	AG(31)	10
8	bnlg2328	AGCAGTGAGGAAGAAGCAGG	TTACCCTCCCTTGTCTGTGAC	AG(33)	2
9	bnlg1812	CGAGAAGACTTGCGTGAACA	TTACGTGCGTCGTGAGAATC	AG(22)	8
10	umc2238	CACCTGCTCAAGCACCATCC	TCATGGAGTACCACCATTGTATCC	(ATGG)4	1
11	umc2239	ACTAGCACTGTGATGAGTGCAG	TTGTACCTCTCCTACGTACCCCA	(TAA)4	1
12	phi057	CTCATCAGTGCCGTCGTCCAT	CAGTCGCAAGAAACCGTTGCC	GCC	7
13	phi112	TGCCCTGCAGGTTACATTGAGT	AGGAGTACGCTTGGATGCTCTTC	AG	7
14	umc1066	ATGGAGCACGTCATCTCAATGG	AGCAGCAGCAACGTCATGACACT	(GCCAGA)5	7
15	phi028	TCTCGCTGCTTCGATTAGTACGG	AATGCAGGCGATGGTTCTCCGGCCT	GAA	9
16	phi033	ATCGAAATGCAGGCGATGGTTCTC	ATCGAGATGTTCTACGCCCTGAAGT	AGG	9
17	phi044	TTATTGGTCCCTCTCCCGTCCAGA	AGCATACCCCAATGGTCAACAGGGA	CCCT	9
18	phi095	CCGATCGGCTTTATCACTGTTTAGC	ATGCACCATTCTAGCACTATAGCAACT	AG	1
19	umc1776	AAGGCTCGTGGCATACTGTAGT	GCTGTACGTACGGGTGCAATG	(TA)7	2
20	nc009	CGAAAGTCGATCGAGAGACC	CCTCTTTCACCCCTTCCTT	AG	6
21	nc010	TGAGCTGACGACGAGCAG	CATTATCTGTTCCGCCCG	GTAC	6
22	phi031	GCAACAGGTTACATGAGCTGACGA	CCAGCGTGTGTTCCAGTAGTT	GTAC	6
23	umc1014	GAAAGTCGATCGAGAGACCCTG	CCCTCTTTCACCCCTTCCTT	(GA)12	6
24	bmc1136	TAACCGGATGAGCATCTTCC	CATCAGCTTCAACGATTCG	(AG)14	6
25	phi017	CGTTGGCGACCAGGGTGCCTGGAT	TGCAACAGCCATTGATCATCAAAC	TAC	9
26	umc1082	CCGACCATGCATAAGGTCTAGG	GCCTGCATAGAGAGTGGTATGAT	(GA)16	1
27	phi022	TGCGCACCAGCGACTGACC	GCGGGCGACGCTTCCAAAC	GTGC	9
28	phi027	CACAGCACGTTGCGGATTTCTCT	GCGTACGTACGACGAAGACAC	GCGCT	9
29	phi061	GACGTAAGCTAGCTCTGCCAT	AAACAAGAACGGCGGTGCTGATTC	TTCT-GTAT	9
30	p-umc2075	CCTGGTATCTTGATGAGCTGGATT	AGAAACATGTCCGGTTCTTCAGAC	(AGCCAG)4	8
31	umc1448	ATCCTCTCATCTTTAGTCCACCG	CATATACAGTCTTCTGCGCTGCTCA	(GCT)5	2
32	umc1080	GAGGAGAAAAGGAGATGAAAAGC	AGATGCCGAGAAGATTCTAAACA	(GA)15	2
33	umc2095	GTGCTGGTTGTGCATGGTATATTG	CCATGCATGACGCAAGTCTCT	(AG)7	9
34	umc1678	GTAGAGATCGATTGCTAACCTGC	AGTTGTTCCGTTCCGCTCTTATC	(TCG)6	10
35	umc1084	GATAAAAAGGCAAGTGAACAAGG	ATATCAACCAGAGGCTGGAACCTTG	(CT)23	10
36	umc2317	ATGTACTGCTCGGCATGGAT	CTGTGGTCGTCTTCTTGGTCG	(GGC)6	6

Tabela S1 - Marcadores SSR selecionados no MaizeGBD

Numero	Nome do Loco	Primer F	Primer R	Tipo de Repetição	Cromossomo
37	p-umc2501	GCCGTCGGTATCCACAAGAG	GCAGCTAGCAGTGTCTTAAACCGT	GAG	6
38	p-umc2569	GTGACACCCTAGCCCTCTTAGACA	TAGCTGGAGTATGTCGCTTGGTG	AGGC	1
39	p-mmc0371	CCACGCACCTCTTGTAAAC	GGGAACCTACAGCTTGGT	(GA)4N1(GA)23(GGA)1(GAA)2	4
40	phi047	GGAGATGCTCGCACTGTTCTC	CTCCACCCTCTTTGACATGGTATG	ATC	3
41	p-umc1024	CCTTTTTCGCCTCGCTTTTTAT	TCGTCGCTCCAATCATACTG	(GA)8	2
42	p-umc1035	CTGGCATGATCAGCTATGTATG	TAACATCAGCAGGTTTGCTATTC	(CT)19	1
43	p-umc1515	AGAGAGGCTGCTCAATAAGTTGC	TTAGTAGTTTCGGTGTCCGTTTCC	(GA)6	1
44	p-umc1358	AGAACCTCCCGCTTGACGAC	ACCTCAACCTCGACCTTGCAT	(CCT)6	1
45	umc1144	ATGGCCCACTCATATCTCTGT	TGTGTTGATTAGCAGCGGATAAAA	(CT)8	1
46	p-umc1676	AGTCGTACGATGACGGAGGC	GCACCACCGACTGATCAAGA	(GCC)7	1
47	umc1503	TTCATGACACACAAACCACAGATG	GCACCCTAGCAGACTACAACATCC	(GAC)4	4
48	p-umc1298	AGCTGAACAAAATAAACGGAACGA	AGGACAAGAAAAGAAGAAGCACG	(GCT)4	1
49	p-phi070	GCTGAGCGATCAGTTCATCCAG	CCATGGCAGGGTCTCTCAAG	AGCTG	6
50	p-umc1063	AGGCCACTGAGCAGGTGAAG	GTGATGGTAGAGGAGTCTTGGTG	(CCAACA)4	6
51	p-umc1914	CAACATGAGCGTCTAAATACTCG	ACAGGAACACATGAGGTCATCAA	(CAG)4	1
52	p-umc1265	GCCTAGTCGCCTACCTACCAAT	TGTGTTCTGATTGGGTGAGACAT	(TCAC)4	2
53	p-umc1542	TAAAGCTATGATGGCACTTGCAGA	CATATTTGCCTTTGCCCTTTTGA	(AG)10	2
54	p-umc2256	GGTCTAGTCGTTAATTCTTTAGCG	GGTCAAGGACTCTTCTTCCTCCTT	(AG)6	3
55	p-bnlgl208	GCTGTGATGGTGAGACGAGA	GCAGGCACTACTAAAACCGC	AG(10)	5
6	p-umc1340	GATGTCTCTATGGAACCCAGCAAC	GAGACGCCTACGAGTACCACAAC	(GCC)7	8
57	p-umc2075	CCTGGTATCTTGATGAGCTGGATT	AGAAACATGTCCGGTCTTCAGAC	(AGCCAG)4	8
58	p-umc1588	TGACAACAGCTATGTGCTGCTCC	GGATGAAGCAAACCAAGCACATAC	(AT)6	9
59	p-umc2287	CTAGCTAGTAACAGAGCATCGCGG	CTGAGGTGTAGGATCGAGCAAGTT	(CT)7	4
60	p-umc1328	TACAAGGAGGAGGCCGCTGT	ATCCAGTCTCCGACTTCCAAC	(TGC)5	4
61	p-umc1796	CGCTGAGGCTTAAGATGGTGTT	AACGCCCTTACGAGCACGAAC	(CCT)5	6
62	p-umc1537	CATGAATCACACTTGGATGTGGTC	AGAAGCTGTCTCGTTCAAGCTC	(TCG)4	5
63	p-umc1370	GGGAGCACACAGTAGTACTCGAT	AGAGGCTCTCCTCTTCAAGCTC	(CGGG)5	9
64	p-umc2122	TTGACAAGCTAGTGTGCAACTGTG	TGAAAGCCCACTGGACAACTAAT	(TG)8	10
65	p-bnlgl601	CAGACCAGAGACCATCTGCA	ATCGTGCGCTAGTCCAGAGT	AG(24)	3
66	p-umc1540	GCGATGATGATGATGATTTATAGCAG	CACTGGAGTCATCTTCTCAGACA	(TGA)6	3
67	p-umc1920	GGTTCGGGTTTGCTACGTGTT	ACGAGACAACAACCAAGACAAA	(GT)8	3
68	p-phi102228	ATCCGACGCAATCAACA	TTCATCTCTCCAGGACCTT	AAGC	3
69	p-nc030	CCCCTGTCTTTCTCTCTCC	CGATTAGATTGGGGTGCG	CT	3
70	p-mmc0022	AGGTGTGTTTTTGTTCGCT	TGCTTGTTAAGCTCATTATT	(CA)6(CT)2(CA)6	3
72	p-phi96100	AGGAGGACCCCAACTCCTG	TTGCACGAGCCATCGTAT	ACCT	2
73	p-umc1560	CGTTCGTCTCTGGTAGCGTAG	TATAACAGCCTGTGCTGCTTG	(GC)6	2
74	p-umc1404	GTTGGCCTTCTCTCTACCCC	CCAGTTCATCAGGTCATCAACCTT	(CCG)4	3
75	p-umc1757	TTTTCTGCAGGATAACATTGTG	ATAGGAGGTGAGGTGAGGAGGAAG	(TCC)7	4
76	p-nc004	TGCGAAGAAGCAGTAGCAA	TGGAGGTAGAAGACGCACG	AG	4
77	p-umc1117	AATTCTAGTCTGGGTCGGAACCTC	CGTGCCGTTGGAGTCTACTACT	(TCGA)4	4

Tabela S1 - Marcadores SSR selecionados no MaizeGBD

Numero	Nome do Loco	Primer F	Primer R	Tipo de Repetição	Cromossomo
78	p-phi079	TGGTGCTCGTTGCCAAATCTACGA	GCAGTGGTGGTTTCGAACAGACAA	AGATG	4
79	p-umc1476	CTCTGCCTCAGTCTGGTCGC	CGAGGAAAGGAAGGAGAGCG	(CGG)5	4
80	p-phi093	AGTGCCTCAGCTTCATCGCTACAAG	AGGCCATGCATGCTTGCAACAATGGATACA	AGCT	4
81	p-umc1423	TAGTATGGTCCATTGATGCTGGC	GAGCAGGCGGAGGATACTAGC	(CTAG)4	5
82	p-umc1679	CACTGCTAAGCTGCTCCCTGTT	TGCTAACTAACCTTGACCCTCTCA	(AAG)5	5
83	p-phi109188	AAGCTCAGAAGCCGGAGC	GGTCATCAAGCTCTCTGATCG	AAAG	5
84	p-mmc0081	TGAAATAATTCACAGCACTCC	TGATAGCACAAACACAGCTATG	(GT)17((GA)8	5
85	p-umc1752	ATCCTCCTCCATATTCTATCGCGT	GAAACAGAGCAGGAACCCGGAG	(CGG)5	5
86	p-umc1792	CATGGGACAGCAAGAGACACAG	ACCTTCATCACCTGCAACTACGAC	(CGG)5	5
87	p-umc1153	CAGCATCTATAGCTTGCATT	TGGGTTTGTGTTGTTGTTGTTG	(TCA)4	5
88	p-umc1401	CTCTGGTCCATCCTCATCGACT	TCTCTTGATCACATATCGATCCCA	(CCA)4	7
89	p-umc1660	GAGCATCGCAGTGGTGGTTAAG	AGTAGAGGGTGAGCGACTTGGAC	(ACG)6	7
90	p-umc1768	ATTCCATTTCCCGAGCCTACTA	CAGAGCACACACTCACGAGTCAC	(CG)8	7
91	p-umc1760	CATGTGGGTGTTAATAAGCAAGGG	GCCTTGTGCTCTCTGAAACAAT	(GA)6	7
92	p-phi116	GCATACGGCCATGGATGGGA	TCCCTGCCGGGACTCCTG	ACTG/ACG***	7
93	p-umc1817	CTACGCAGGCTTCAACCACC	GTAAGTGGTATGATGGTACCCTG	(CAG)6	8
94	p-umc1304	CATGCAGCTTCCAAATTAATCC	GCCAACTAGAAGTACTGCTGCTCC	(TCGA)4	8
95	p-umc1460	GCTCAATCGTAGTAACAGCAGCAG	TCTGCAC TAGAATGGCTTGGTACA	(GCA)4	8
96	p-umc1728	AGTACTTTCAGGCAGGGACCTTCT	AACGCCTTCTTGTAGCTGTAGGG	(GAC)5	8
97	p-umc1152	CCGAAGATAACCAACAATAATAGTAGG	ACTGTACGCCTCCCTTCTC	(ATAG)6	10
98	p-umc1582	AGATTACGTAGCCACGCTTATTCG	GTGCGTGTGAGAGTGATATCGAG	(TACG)4	10
99	p-umc1367	TGGACGATCTGCTTCTCAGG	GAAGGCTTCTTCCCGAGTAGGTC	(CGA)6	10
100	p-umc1196	CGTGCTACTACTGCTACAAAGCGA	AGTCGTTCTGTCTTCCGAAACT	CACACG	10
101	p-umc1412	GCATCTGTAGCCTTTTGTGTGTG	CTCAGCTTGACGTTATCGCTT	(CCA)5	7
102	p-phi034	TAGCGACAGGATGGCCTCTTCT	GGGGAGCACGCCTTCTTCT	CCT	7
103	p-umc1607	ACTAATTTTCGGTAGTCGTGTCG	GGAAAGAGAGAGGCTGTAGGTGGT	(TGC)5	8
104	p-phi080	CACCCGATGCAACTTGCCTAGA	TCGTACGTTCCACGACATCAC	AGGAG	8
105	p-umc1576	TGTACAAAATTACAAGTGGCAGC	CTCGTCATCCTTTCTGAGTGTAT	(TG)6	10
106	p-umc1239	ATCAACACACCTTTCGATTTCTGG	CGGTGATTAGTCGATGAAGAGTGA	(TG)6	10

Figura S1. Parâmetros de busca na ferramenta BLAST do MaizeGDB otimizados para seqüências curtas (Short sequences)

MaizeGDB BLAST

Please use the NCBI BLAST tool to BLAST against GenBank targets.
BLAST up to 5 sequences of up to 35,000 bp total length against maize datasets.
 Other BLAST pages: [Gramene.org](#), [PlantGDB](#), [NCBI](#),

Step 1: input your sequences (Raw sequence, FASTA, or Genbank accessions) ?

Sequence type: Nucleotides Amino Acids
 Enter your sequence: [\(example\)](#)
 GCTGTGATGGTGAGACGAGA

or upload from file: Sin archivos seleccionados

Step 2: select datasets ?

Filter targets by
 Include: genomes gene models both
 Select a BLAST target
 Zm-B73-REFERENCE-NAM-5.0

Step 3: select BLAST parameters ?

Optimize for: High similarity Low similarity Short sequences
 advanced settings

Step 4: select output type ?

Enhanced output BLAST table output BLAST text output

Name your sequence (optional) C3F ?
 E-mail (optional) ?

Citation: Cannon et al., 2011 "POPCorn: An Online Resource Providing Access to Distributed and Diverse Maize Project Data"

Contact Please cite us!

Fonte: MaizeGDB