

UNIVERSIDADE FEDERAL DE VIÇOSA

Importância de características auxiliares no melhoramento da produção e da capacidade de expansão de milho-pipoca com abordagem de inteligência computacional

Kellton de Oliveira Sabino
Magister Scientiae

**VIÇOSA - MINAS GERAIS
2025**

KELLTON DE OLIVEIRA SABINO

Importância de características auxiliares no melhoramento da produção e da capacidade de expansão de milho-pipoca com abordagem de inteligência computacional

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

Orientador: Cosme Damiao Cruz

**VIÇOSA - MINAS GERAIS
2025**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

S116i
2025 Sabino, Kellton de Oliveira, 1992-
Importância de características auxiliares no melhoramento da produção e da capacidade de expansão de milho-pipoca : com abordagem de inteligência computacional / Kellton de Oliveira Sabino. – Viçosa, MG, 2025.

1 dissertação eletrônica (97 f.): il. (algumas color.).

Inclui anexos.

Orientador: Cosme Damião Cruz.

Dissertação (mestrado) - Universidade Federal de Viçosa, Departamento de Estatística, 2025.

Referências bibliográficas: f. 77-81.

DOI: <https://doi.org/10.47328/ufvbbt.2025.718>

Modo de acesso: World Wide Web.

1. Decisão estatística. 2. Milho pipoca - Rendimento - Métodos estatísticos. 3. Topologia. I. Cruz, Cosme Damião, 1958-. II. Universidade Federal de Viçosa. Departamento de Estatística. Programa de Pós-Graduação em Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 519.55

KELLTON DE OLIVEIRA SABINO

Importância de características auxiliares no melhoramento da produção e da capacidade de expansão de milho-pipoca com abordagem de inteligência computacional

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 4 de agosto de 2025.

Assentimento:

Kellton de Oliveira Sabino
Autor

Cosme Damiao Cruz
Orientador

Essa dissertação foi assinada digitalmente pelo autor em 31/10/2025 às 12:25:23 e pelo orientador em 31/10/2025 às 18:18:41. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **MS7Y.J6JE.DFZX** e clique no botão 'Validar documento'.

Aos meus pais e irmãos

AGRADECIMENTOS

Quero agradecer, primeiramente, a Deus, por me permitir chegar até aqui com saúde e força.

Não sei o que seria de mim sem minha mãe e o Nem, meus pais, que me guiaram e educaram. Esta dissertação não poderia ser dedicada a ninguém além de vocês. Obrigado pelo incansável esforço e amor. Deu certo, mãe! Esta conquista é nossa. Obrigado, Nem, por ser um pai zeloso. Eu devo isso a você. Muito obrigado! Não tenho palavras para descrever o amor e a admiração que sinto por vocês.

Agradeço também ao meu irmão William, que foi pioneiro em tudo isso. Que orgulho eu tenho de você! Obrigado por abrir caminhos e mostrar que podemos ser donos do nosso futuro. Sua garra e determinação me inspiram. Quando eu crescer, quero ser igual a você.

Agradeço ao meu irmão Pedrinho, meu bebê, minha dupla, a quem eu amo de paixão. Obrigado por me aguentar, por estar comigo e por cuidar de mim. Você é meu mundo.

À minha tia Márcia, que sempre zelou pela minha educação e nunca duvidou de minha capacidade, obrigado pelos puxões de orelha e pelos ensinamentos de vida. Ao meu namorado Kike, parceiro de vida, pela calma e paciência no dia a dia. Obrigado por me acalmar, por me entender e por me amar. Com você, meus dias são mais felizes.

Aos amigos que Viçosa me deu a honra de conhecer, meus MatMigos: Rodrigo, Sara, Brenda, Bernardo, Carol, Fernanda, Luisa, Isabela, Letícia e Mari. Obrigado por dividirem a jornada comigo e tornarem minha vida mais bela. Obrigado, Letícia, pelas horas de conversa e estudo.

Um agradecimento especial à minha amiga Mari, por fazer minha inscrição no mestrado, por cuidar de mim e por ser mais que uma amiga: uma irmã tão cuidadosa e amável. Estaremos juntos para sempre. Obrigado por todo o amor e cuidado.

Aos amigos que nunca esquecerei, por serem tão diferentes ou tão parecidos comigo, mas que dividem a caminhada dia após dia. Obrigado, Pamela, pelos tantos momentos de conversa na Rita, por me fazer admirar você tanto — você é incrível; à minha amiga Luisa, a quem levo no coração, obrigado pelos melhores momentos de risada, "eu facilmente..". Ao meu grande amigo João, obrigado pela simplicidade de uma amizade sincera e por todas as loucuras da vida.

À Universidade Federal de Viçosa, por ser minha casa por tantos anos, pelo curso de Matemática e, agora, pela oportunidade de realizar a pós-graduação. Ao PPESTBIO, pela oportunidade de estudo, acolhida e apoio.

Agradeço a cada um dos professores que me permitiram chegar até aqui. Obrigado por me ajudarem ao longo dessa caminhada. Vocês são demais!

Agradeço ao meu orientador, Cosme. Obrigado por sempre me apoiar e ensinar, por agregar tanto à minha vida e por me orientar com dedicação.

Agradeço aos professores José Ivo e Amaral por participarem da minha banca. Com certeza, a contribuição de vocês será de grande valor para este trabalho.

Agradeço às agências de fomento: ao Carrefour, pela concessão da bolsa de estudos, e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Quero agradecer a CAPES. Por fim, gostaria de agradecer a todos que, de forma direta ou indireta, me ajudaram ou torceram para que eu pudesse alcançar mais essa vitória. Meu muito obrigado!

RESUMO

SABINO, Kellton de Oliveira, M.Sc., Universidade Federal de Viçosa, agosto de 2025. **Importância de características auxiliares no melhoramento da produção e da capacidade de expansão de milho-pipoca com abordagem de inteligência computacional.** Orientador: Cosme Damiao Cruz.

O presente estudo avalia o desempenho de linhagens de milho-pipoca (*Zea mays* L. var. *everta*) sob condições contrastantes de disponibilidade hídrica, com o objetivo de identificar características auxiliares associadas ao rendimento de grãos e à capacidade de expansão. Os ensaios foram conduzidos nas safras de 2020 e 2021, em delineamento em blocos ao acaso, considerando manejo bem irrigado e sob estresse hídrico. Foram avaliados caracteres agronômicos, fisiológicos e morfológicos do sistema radicular, e aplicadas análises de variabilidade genética, coeficientes de determinação e correlações fenotípicas, genotípicas e parciais para embasar a interpretação dos dados. Tradicionalmente, a análise de trilha tem sido empregada no melhoramento de plantas para decompor correlações em efeitos diretos e indiretos. Contudo, este trabalho apresenta um outro método para o mesmo fim: o uso de árvores de decisão. Por meio da construção de árvores de regressão baseadas no algoritmo CART, a pesquisa utiliza a topologia da árvore — formada pelos nós e divisões hierárquicas — para evidenciar quais variáveis explicativas exercem maior influência sobre as variáveis principais. Essa abordagem permite lidar com relações não lineares e captar interações complexas entre caracteres, oferecendo uma visão prática e interpretável da relevância relativa de cada variável. Os resultados demonstram variabilidade genética significativa entre as linhagens estudadas e mostram que a estrutura hierárquica das árvores de decisão permite identificar, de forma objetiva, os caracteres auxiliares mais impactantes sobre o rendimento de grãos e a capacidade de expansão. Dessa forma, o estudo amplia as possibilidades analíticas disponíveis para programas de melhoramento de milho-pipoca, fornecendo uma alternativa sólida à análise de trilha.

Palavras-chave: milhopipoca; árvores de decisão; topologia.

ABSTRACT

SABINO, Kellton de Oliveira, M.Sc., Universidade Federal de Viçosa, August, 2025. **Importance of auxiliary characteristics in improving the yield and expansion capacity of popcorn using a computational intelligence approach.** Adviser: Cosme Damiao Cruz.

This study evaluates the performance of popcorn maize (*Zea mays* L. var. *everta*) lines under contrasting water availability conditions, aiming to identify auxiliary traits associated with grain yield and popping expansion. The experiments were carried out during the 2020 and 2021 growing seasons in a randomized block design, considering both well-irrigated management and water stress conditions. Agronomic, physiological, and root system morphological traits were assessed, and analyses of genetic variability, coefficients of determination, and phenotypic, genotypic, and partial correlations were performed to support data interpretation. Traditionally, path analysis has been employed in plant breeding to decompose correlations into direct and indirect effects. However, this study presents another method for the same purpose: the use of decision trees. By constructing regression trees based on the CART algorithm, the research uses the tree topology—formed by nodes and hierarchical splits—to highlight which explanatory variables exert the greatest influence on the main traits. This approach handles nonlinear relationships and captures complex interactions among traits, offering a practical and interpretable view of each variable's relative importance. The results demonstrate significant genetic variability among the studied lines and show that the hierarchical structure of decision trees makes it possible to objectively identify the most impactful auxiliary traits on grain yield and popping expansion. Therefore, this study broadens the analytical options available for popcorn maize breeding programs, providing a robust alternative to path analysis.

Keywords: popcorn maize; decision trees; topology

SUMÁRIO

1. INTRODUÇÃO.....	10
2. REVISÃO DE LITERATURA.....	11
2.1. Melhoramento Genético de Milho-Pipoca.....	11
2.2. Abordagens biométricas para identificação de caracteres auxiliares eficazes	15
2.2.1. Análise de trilha como método de identificação de características auxiliares.....	15
2.2.2. Métodos de árvores de decisão em melhoramento genético.....	17
3. MATERIAIS E MÉTODOS.....	21
3.1 Material Experimental	21
3.2 Potencial das linhagens de milho pipoca e quantificação da variabilidade genética	25
3.3 Relações entre características auxiliares e a produção e a capacidade expansão de milho pipoca	26
3.4 Identificação de caracteres auxiliar no melhoramento da produção e capacidade de expansão do milho pipoca.....	28
3.4.1 Abordagem fundamentada em associação linear	28
3.4.2 Abordagem fundamentada em associação não-linear.....	31
4. RESULTADOS E DISCUSSÃO	33
4.1 Potencial de genótipos de milho pipoca e variabilidade.....	33
4.2 Associações entre caracteres agrônômicos auxiliares e a capacidade de expansão e rendimento de grãos.....	43
4.3 Estudo dos efeitos totais, diretos e indiretos de caracteres auxiliares sobre a capacidade de expansão e rendimento de grãos: Análise de trilha	51
4.4 Estudo dos efeitos totais, diretos e indiretos de caracteres auxiliares sobre a capacidade de expansão e rendimento de grãos: Árvore de decisão.....	63
4.4.1. Eficiência de predição de características principais por meio de características auxiliares.....	64
4.4.2. Árvores de decisão.....	65
4.4.3. Comparação.....	72
5. CONCLUSÕES	76

REFERÊNCIAS	77
APÊNDICE A	82
APÊNDICE B.....	90

1. INTRODUÇÃO

O milho-pipoca (*Zea mays* L. var. *everta*) é uma variedade especial de milho que chama atenção pela sua capacidade única de expansão, uma característica que vai além do aspecto comercial e envolve também questões culturais e sociais (MENEGALDO, 2011; COSTA, 2022). No Brasil, esse grão tem presença marcante em festas populares, sessões de cinema e no dia a dia das famílias (PACHECO et al., 1996). Apesar de sua popularidade, o cultivo do milho-pipoca ainda enfrenta muitos desafios — e entre eles, o estresse hídrico é um dos mais preocupantes. Por depender de uma boa distribuição de chuvas ao longo do ciclo, essa cultura pode sofrer grandes perdas em produtividade e qualidade dos grãos quando exposta a condições ambientais desfavoráveis (GAMA et al., 1990; DALLACORT et al., 2011; IPCC, 2022; LESK et al., 2016).

Nesse cenário, o melhoramento genético surge como uma ferramenta essencial. O objetivo é claro: desenvolver plantas mais produtivas, com grãos de melhor qualidade e, principalmente, mais resistentes à seca (SAWAZAKI, 1996; CARANHATO et al., 2022). Mas para que isso aconteça de forma eficiente, é preciso entender melhor como diferentes características da planta — como altura, estrutura radicular, acúmulo de pigmentos e outros fatores agrônômicos e fisiológicos — se relacionam com o desempenho em ambientes com e sem estresse hídrico (RIBEIRO et al., 2024; KAMPHORST et al., 2021; GAO & LYNCH, 2016).

Durante muito tempo, análises estatísticas clássicas, como a análise de trilha, têm sido usadas para investigar essas relações (RANGEL et al., 2011; CRUZ et al., 2012; BASU et al., 2017). Essa técnica é importante porque permite identificar o efeito direto e indireto da relação entre uma característica e outra. No entanto, por depender de relações lineares, pode-se mostrar limitada, não identificando padrões complexos, comuns em dados de melhoramento genético (GARCIA-CAMPOS et al., 2015; KLEM, 1995).

Diante dessas limitações, surge a necessidade de considerar metodologias analíticas mais eficazes e capazes de lidar com tais complexidades. As árvores de decisão, por exemplo, embora tenham surgido com foco em classificação e predição, podem ser adaptadas para ajudar no entendimento de como características auxiliares se organizam e influenciam diretamente variáveis principais, como o rendimento de grãos e a capacidade de expansão (DE'ATH & FABRICIUS, 2000; IZEMAN, 2013; CHATTERJEE & SIMONOFF, 2022). A estrutura gerada por essas árvores, de forma visual e hierárquica, se assemelha ao diagrama de causa e efeito da análise de trilha — com a vantagem de lidar com relações não lineares e

interações entre múltiplas variáveis (PHALAK et al., 2014; THOMAS et al., 2019).

Assim, ao explorar e comparar essas metodologias aplicadas a dados agronômicos, fisiológicos e morfológicos, este estudo busca indicar caminhos mais precisos para a seleção de genótipos superiores, especialmente frente aos desafios do estresse hídrico (RINALDI et al., 2022; WASHBURN et al., 2021). Espera-se, com isso, contribuir para o avanço do melhoramento do milho-pipoca no Brasil, fortalecendo uma cultura tão presente no cotidiano da população (CARANHATO et al., 2022; SAWAZAKI, 1996).

2. REVISÃO DE LITERATURA

2.1 Melhoramento Genético de Milho-Pipoca

- **Histórico e Avanços Recentes**

O milho é o cereal mais cultivado no mundo, dado o seu baixo valor e sua potente integração em diferentes sistemas de produção. Ele é essencial tanto na agricultura intensiva, voltada para a produção em larga escala, quanto na agricultura familiar, que muitas vezes constitui a base do sistema agrícola. Além de servir como alimento para humanos e animais, ele é utilizado diretamente na propriedade, com o excedente sendo comercializado ou trocado por insumos e outros produtos, especialmente através de cooperativas, que oferecem maior segurança e praticidade (MACHADO et al., 2014).

Na culinária brasileira, o milho tem presença marcante em todas as regiões, sendo uma importante fonte de energia para pessoas e animais, Menegaldo (2011). O milho-pipoca, considerado um tipo diferenciado, destaca-se por suas características únicas, que tem sua diferenciação do milho comum dada pela sua capacidade de expansão, tem uma maior peculiaridade sendo servida para o consumo humano. O Brasil ocupa o segundo lugar no ranking mundial de produção de milho-pipoca, um alimento amplamente consumido em diversos contextos, como em festas populares — a exemplo das festas juninas — no cinema e no ambiente doméstico. Essa cultura possui grande relevância econômica, devido ao seu importante valor comercial, uma vez que é comercializada a preços superiores aos do milho convencional. Cabe destacar, ainda, que a pipoca é um alimento altamente nutritivo, rico em antioxidantes, proteínas, fibras, vitaminas, açúcares e gorduras, o que contribui para seu destaque como uma opção saudável e saudável (COSTA, 2022).

O rendimento de grãos e a capacidade de expansão são os principais atributos de interesse no melhoramento do milho-pipoca, atendendo tanto aos produtores, que buscam alta

produtividade, quanto aos consumidores, que valorizam a qualidade da pipoca, como textura e maciez dos flocos (SAWAZAKI, 1996). A capacidade de expansão, característica que define o volume da pipoca em relação à massa dos grãos antes do estouro (PACHECO et al., 1996), é fundamental para o valor comercial dessa cultura, superando o do milho comum. No entanto, os cuidados durante a colheita e o armazenamento são cruciais para evitar danos ao pericarpo, o que poderia comprometer a pressão interna necessária para a expansão ideal dos grãos (SAWAZAKI, 1996).

O aumento da produção e do consumo de milho-pipoca tornam ainda mais relevantes estudos que relacionam a composição desse produto aos benefícios que o mesmo pode fornecer a saúde de quem o consome.

O melhoramento genético do *Zea mays everta* tem uma longa história, começando com as civilizações nativas do continente americano (do Norte e Central), que foram as primeiras a cultivar e domesticar variedades de milho, incluindo aquelas que apresentam traços de expansão quando expostas ao calor (GAMA et al., 1990). No decorrer dos séculos, o milho-pipoca evoluiu para se tornar uma cultura economicamente importante, com foco não apenas em sua capacidade de expansão, mas também em características como resistência a doenças e adaptação a diferentes condições ambientais (SAWAZAKI, 1996).

Segundo Gama et al. (1990), o sistema de cultivo do milho-pipoca requer boa supervisão ao longo de todo o ciclo, especialmente nas fases críticas de pendramento e enchimento de grãos, sendo sensíveis a déficits hídricos e temperaturas extremas. Idealmente, a oferta deve ser em torno de 600 mm, com boa distribuição ao longo do ciclo, evitando períodos em que a água disponível no solo seja inferior a 40%. A temperatura ideal para o cultivo está em torno de 30°C, com limites entre 10°C e 40°C. Outros fatores, como radiação, vento e umidade, também influenciam indiretamente, principalmente por meio da evapotranspiração e aumento.

Segundo Dallacort et al. (2011), a distribuição das chuvas tem influência direta no planejamento agrícola, afetando o desenvolvimento das culturas, a produtividade e o manejo dos sistemas hídricos.

Conforme destacado pelo IPCC (2022), as variações do clima, intensificadas pelo aquecimento causado pelas atividades humanas, representam uma ameaça crescente à humanidade e à vida no planeta. Dentre os impactos, a seca se destaca como o estresse abiótico mais complexo e prejudicial às culturas agrícolas, especialmente ao milho, que depende significativamente de recursos hídricos para seu desenvolvimento pleno (LESK et al., 2016). O fornecimento adequado de água é crucial, principalmente nas fases de

florescimento e enchimento de grãos, condições nas quais o estresse hídrico pode causar perdas expressivas na produtividade. As variações ambientais podem afetar o desempenho produtivo de distintas cultivares, com os híbridos a apresentarem por vezes rendimentos mais baixos em condições desfavoráveis (SAKIN et al., 2005).

Dessa forma, a busca por características que auxiliem no melhoramento de milho-pipoca e o emprego de métodos que permitam identificar estas características mais promissoras, têm se mostrado de grande potencial para otimizar a produção e a qualidade dos grãos (medido pela sua capacidade de expansão) sob essas condições adversas. Visto que esse tipo de cultivar exige boa distribuição de chuva durante todo o ciclo. Os programas de melhoramento visam desenvolver híbridos com alta expansão de estalo (40 mL g^{-1}) e produtividade (4.000 kg ha^{-1}) (CARANHATO et al., 2022).

- **Características Auxiliares Importantes no Melhoramento de Milho-Pipoca Tolerantes a Seca**

Desenvolver variedades de milho-pipoca mais tolerantes à seca é um dos grandes desafios atuais da pesquisa agrícola, especialmente frente às mudanças climáticas e à crescente escassez hídrica em diversas regiões produtoras. Essa tolerância, no entanto, não depende de um único fator, mas sim de um conjunto de características auxiliares que atuam de forma integrada para garantir o bom desempenho da planta mesmo sob estresse hídrico.

Dentre essas características, aspectos morfológicos e fisiológicos do sistema radicular ganham destaque. O comprimento das raízes, seu volume, área de superfície e peso seco são fatores determinantes para uma maior eficiência na captação e retenção de água em solos secos (RIBEIRO et al., 2024). Ademais, adaptações na arquitetura radicular, como ângulos mais agudos da raiz principal e da raiz da coroa, permitem que o sistema se aprofunde no solo em busca de umidade (KAMPHORST, 2018). A presença de menor número de raízes da coroa e uma menor densidade de ramificações também contribuem para reduzir o custo metabólico e otimizar a absorção de água.

As características do rebento, como o comprimento e o peso seco, também exercem papel fundamental, ajudando a planta a manter seu crescimento mesmo quando a oferta hídrica está limitada (CARVALHO et al., 2023). Complementando esse cenário, fatores fisiológicos como potencial hídrico foliar, controle osmótico, fechamento estomático e

produção de biomassa também são frequentemente associados à capacidade da planta de tolerar períodos de seca (TINOVONGA et al., 2024).

Outro aspecto relevante são os metabólitos, que participam da regulação das respostas ao estresse. Essas substâncias ajudam não só no entendimento do comportamento da planta frente à seca, mas também no mapeamento da arquitetura genética associada à tolerância, possibilitando avanços na seleção assistida por marcadores moleculares (LI et al., 2021; CARANHATO et al., 2022).

No campo, várias dessas características morfofisiológicas são, relativamente, de mais fácil mensuração e podem ser usadas como critérios indiretos para seleção. O peso de 100 grãos (GW), a prolificidade (PR), o comprimento do pendão (TL) e o ângulo foliar (FA) são exemplos de atributos que apresentam forte correlação com o rendimento de grãos e a capacidade de expansão — duas das principais variáveis econômicas do milho-pipoca. Além disso, o índice de verdor das folhas, medido por ferramentas como SPAD e NDVI, tem se mostrado eficaz para monitorar a saúde da planta durante o enchimento dos grãos, sendo útil para uso como critério de seleção indireta em programas de melhoramento (Kamphorst et al., 2021).

A prolificidade, por exemplo, apresentou a maior associação com a capacidade de expansão em condições de seca, enquanto o índice SPAD, medido 17 dias após o florescimento masculino, destacou-se como uma ferramenta promissora para seleção indireta de genótipos com maior rendimento de grãos (KAMPHORST et al., 2021).

A importância dessas características não está apenas na sua relação direta com o desempenho da planta, mas também na possibilidade de combiná-las com estratégias de melhoramento genético. Com o suporte de análises estatísticas avançadas, como regressão múltipla, é possível identificar quais traços têm maior poder preditivo e utilizá-los de forma eficiente na seleção de materiais genéticos com maior adaptação ao estresse hídrico (KAMPHORST et al., 2021).

Além disso, estudos mais recentes têm mostrado que raízes profundas, maior prolificidade e menor número de ramificações no pendão também contribuem para a adaptação da planta ao déficit (GAO & LYNCH, 2016). Características como essas, quando

integradas com dados genéticos e fenotípicos, formam a base para programas de melhoramento mais robustos e eficientes.

Em suma, entender e explorar essas características auxiliares são estratégias essenciais para o sucesso do melhoramento genético do milho-pipoca. Elas não apenas ajudam a explicar como as plantas reagem ao estresse hídrico, mas também oferecem caminhos práticos para o desenvolvimento de cultivares mais resilientes, contribuindo diretamente para a segurança alimentar em contextos de escassez de água.

2.2. Abordagens Biométricas para Identificação de Caracteres Auxiliares Eficazes

A eficácia de um programa de melhoramento genético depende, em grande medida, da correta identificação de características que estejam associadas de forma direta e significativa ao caráter de interesse. No caso do milho-pipoca, sob condições de estresse hídrico, a produtividade de grãos e a capacidade de expansão tendem a sofrer reduções consideráveis na herdabilidade e alta influência ambiental. Nesse sentido, a utilização de características auxiliares, também conhecidas como caracteres secundários, torna-se uma estratégia relevante para otimizar a seleção indireta.

A fim de que tais características sejam efetivamente incorporadas aos critérios de seleção, é necessário o uso de abordagens biométricas que permitam avaliar suas relações com o caráter principal. Dentre os métodos disponíveis, destacam-se a análise de trilha, pela sua capacidade de decompor correlações em efeitos diretos e indiretos, e os métodos de árvores de decisão, que integram princípios de aprendizado de máquina para classificar, hierarquizar e prever comportamentos genotípicos com base em múltiplas variáveis. A seguir, são apresentados os fundamentos teóricos e aplicações desses dois métodos no contexto do melhoramento genético de milho-pipoca tolerante à seca.

2.2.1. Análise de Trilha como Método de Identificação de Características Auxiliares

A análise de trilha tem uma origem marcante na história da estatística genética, tendo sido introduzida por Sewall Wright em 1921. À época, a proposta de usar correlações para descrever relações causais entre características ainda causava estranhamento no meio científico. Com o passar dos anos, no entanto, esse método foi sendo aprimorado e acabou se consolidando como uma ferramenta central no desenvolvimento da modelagem por equações

estruturais (MULAİK, 2014). A proposta de Wright foi particularmente inovadora ao permitir que os coeficientes de correlação fossem decompostos em efeitos diretos e indiretos, oferecendo uma nova lente para entender os mecanismos causais entre variáveis (G. et al., 2024).

Devido à exigência da formulação prévia de modelos teóricos que especifiquem relações ao relacionar as variáveis, a análise de trilha depende fortemente do conhecimento técnico e empírico do pesquisador. Ela não tem por objetivo estabelecer causalidade definitiva, mas sim verificar se os dados observados conforme a estrutura de relações proposta (“Path analysis”, 2022; STREINER, 2005). Essa característica torna o método valioso para estudos exploratórios e confirmatórios, especialmente em áreas que envolvem múltiplos traços interdependentes, como o melhoramento genético de plantas.

No campo da genética vegetal, a análise de trilha tem sido amplamente utilizada para investigar como características secundárias se relacionam com traços de interesse agrônomo, como o rendimento de grãos e a capacidade de expansão em milho-pipoca (RANGEL et al., 2011). A técnica permite ao pesquisador entender quais dessas variáveis têm influência direta no caráter principal, favorecendo decisões mais precisas na seleção. Através de uma série de regressões lineares simultâneas, é possível quantificar o peso de cada variável no resultado — uma informação crucial em programas de melhoramento.

Segundo Boker e McArdle (2014), a análise de trilha pode ser representada graficamente por diagramas que ilustram as relações entre as variáveis do modelo. Esses esquemas visuais ajudam a interpretar as interações e permitem uma compreensão mais intuitiva dos efeitos diretos e indiretos que se estabelecem no sistema. Essa estrutura também facilita a estimativa da matriz de momentos esperada, contribuindo para avaliar a adequação do modelo teórico aos dados empíricos.

As correlações fenotípicas ou genotípicas, entre as variáveis são decompostas em efeitos diretos e indiretos, o que ajuda a identificar quais características contribuem diretamente para o desempenho desejado (CRUZ et al., 2012). Na prática, isso significa que a análise de trilha tem sido eficaz, por exemplo, em estudos de milho-pipoca voltados à tolerância à seca, onde características como profundidade radicular e eficiência do uso da água têm sido avaliadas quanto ao impacto direto na produtividade sob estresse hídrico. Da mesma forma, outras características morfofisiológicas, como área foliar e acúmulo de matéria

seca, também têm se mostrado fortemente relacionadas ao rendimento, conforme observado por Basu et al. (2017).

Além disso, a técnica tem revelado possíveis conflitos na seleção de múltiplas características. Um estudo conduzido por Cabral et al. (2021), por exemplo, identificou correlação genotípica negativa entre rendimento e capacidade de expansão, sugerindo que a seleção simultânea de ambos os caracteres pode exigir estratégias mais cuidadosas. Outros trabalhos mostram que traços como o peso da espiga possuem efeito direto positivo sobre a expansão do grão, reforçando seu valor como critério de seleção (VIJAYABHARATHI et al., 2009).

No entanto, apesar de sua utilidade, a análise de trilha tem limitações que precisam ser consideradas. A principal delas está na suposição de linearidade entre as variáveis. Em muitos sistemas biológicos, como os observados em ambientes estressantes ou com múltiplas variáveis em interação, essas relações podem não ser lineares. Isso significa que o método pode falhar ao representar com fidelidade a complexidade dos dados, levando a conclusões imprecisas (GARCIA-CAMPOS et al., 2015). Outro ponto de atenção é a dificuldade de extrapolar os resultados para diferentes ambientes, o que exige validação adicional (KLEM, 1995).

Diante disso, embora a análise de trilha continue sendo uma ferramenta importante no arsenal estatístico dos programas de melhoramento, ela não deve ser utilizada de forma isolada. Métodos mais recentes, como as árvores de decisão e outros algoritmos de aprendizado de máquina, mesmo com o uso não preconizado para este fim, podem ser adaptados para oferecer vantagens importantes quando as relações entre variáveis fogem da linearidade. Assim, a integração entre diferentes abordagens analíticas surge como um caminho promissor para lidar com a complexidade dos dados genéticos, sobretudo em culturas como o milho-pipoca, que exigem eficiência em múltiplas frentes seletivas sob condições ambientais adversas.

2.2.2. Métodos de Árvores de Decisão em Melhoramento Genético

As árvores de decisão são modelos estatísticos amplamente utilizados para fins de resolver problemas relativos a classificação e predição. Deve ser destacado que, em estudos de predição, o objetivo primordial é a busca de procedimentos com melhor desempenho em

métricas tradicionais tais como erros quadráticos e coeficientes de determinação, usando, para tal, refinamentos das árvores de decisão, reduções de dimensionalidade, processos mais robustos de aprendizado, dentre outros. Por outro lado, para fins de estudo de relacionamento entre uma variável principal e outras auxiliares, a teoria de árvore de decisão, focada em princípios de regressão, sem grandes refinamentos, já seria suficiente para auxiliar o melhorista no mapeamento de características auxiliares importantes sobre uma principal, evidenciando seus impactos diretos e suas intrincadas interações de forma que o perfil da árvore seja instrumento valioso para tal fim de interpretação.

As árvores de decisão operam a partir do particionamento recursivo dos dados, um processo que consiste em dividir, sucessivamente, um conjunto de dados em subconjuntos menores com base em critérios definidos. Cada divisão representa uma decisão que leva a diferentes ramificações, construindo uma estrutura hierárquica em forma de árvore. Ao final desse processo, cada caminho que vai da raiz até um nó folha corresponde a uma regra específica de classificação ou previsão. Trata-se de um método não paramétrico, de fácil interpretação e aplicação intuitiva, razão pela qual se consolidou como uma técnica popular nas áreas de aprendizado de máquina e análise estatística (IZENMAN, 2013).

No campo do melhoramento genético de plantas, especialmente quando se deseja explorar a complexidade das características agronômicas, métodos fundamentados em árvores de decisão têm se destacado. Técnicas como Random Forest, Gradient Boosting e outras variações se mostraram eficazes na identificação de características secundárias que influenciam características principais, como o rendimento de grãos. Essas abordagens operam por meio da divisão do espaço fenotípico ou genotípico em regiões que maximizam a homogeneidade interna dos dados, com base em uma variável de interesse, seja ela categórica (em tarefas de classificação) ou contínua (em tarefas de regressão) (PHALAK et al., 2014; THOMAS et al., 2019).

A construção de uma árvore de decisão se inicia a partir de um nó raiz, que representa a totalidade do conjunto de dados. A partir daí, o algoritmo avalia possíveis divisões baseadas nos atributos disponíveis e seleciona aquela que melhor separa os dados. Cada divisão subsequente ocorre de forma recursiva, gerando novos nós até que os dados sejam suficientemente homogêneos ou atendam a critérios de parada pré-definidos. As divisões são guiadas, geralmente, por regras simples — como comparações entre valores de características

e certos limiares —, e continuam até que cada grupo final (ou nó folha) represente uma classificação ou previsão específica (ZHOU, 2022).

Após a construção da árvore, é comum realizar uma etapa de poda, cujo objetivo é evitar o sobreajuste. Nessa fase, removem-se ramificações que contribuem pouco para a generalização do modelo. Técnicas como o Cost Complexity Pruning e o Reduced Error Pruning são frequentemente aplicadas para esse fim (PHALAK et al., 2014). A natureza hierárquica dessas árvores permite não apenas uma modelagem eficiente, mas também uma interpretação clara dos critérios que levaram a determinadas decisões, o que representa uma vantagem importante no contexto de pesquisa aplicada (IZENMAN, 2013). Na proposta deste trabalho considera-se que a poda seja dispensável, tendo em vista que se deseja contemplar o máximo das associações e caminhos que permitem avaliar a relevância do impacto das variáveis principais sobre a variável resposta dadas pela capacidade de expansão e produção do milho pipoca.

Quando aplicadas à regressão, essas estruturas são conhecidas como árvores de regressão e são utilizadas para prever variáveis contínuas, sendo esta situação a mais aplicada aos propósitos deste estudo. Elas dividem o espaço de entrada em regiões e, em cada uma delas, ajustam-se modelos simples para estimar o valor da variável-alvo. Em geral, cada nó folha representa a média dos valores de saída dos dados pertencentes àquela partição (DE'ATH; FABRICIUS, 2000; CHATTERJEE, S.; SIMONOFF, J, 2022).

Dentre as principais vantagens dessas técnicas no contexto biológico e agrônomo, destaca-se sua capacidade de modelar relações não lineares entre variáveis. Isso é especialmente relevante quando se lida com dados biológicos que, frequentemente, apresentam interações complexas que são dificilmente captadas por métodos lineares tradicionais. Além disso, a interpretabilidade desses modelos favorece a compreensão das relações entre características secundárias e as características principais de interesse, como o rendimento de grãos, a tolerância à seca, dentre outros (DE'ATH; FABRICIUS, 2000; CHATTERJEE, S.2000; SIMONOFF, J, 2022).

Outro aspecto importante é a possibilidade de integrar essas técnicas com dados obtidos por sensoriamento remoto, como imagens multiespectrais captadas por VANTs, por exemplo. Em um estudo de Banerjee et al. (2021), árvores de regressão foram utilizadas com sucesso para estimar a emergência de plântulas de trigo a partir desse tipo de dado, o que

evidencia o potencial dessas ferramentas para aplicações em larga escala e com dados complexos.

Para os programas de melhoramento, essa capacidade de identificar os caracteres auxiliares mais influentes permite aos pesquisadores focarem na seleção de características-chave, otimizando tempo, recursos e esforços. Por exemplo, atributos como altura de planta, intensidade de ferrugem e acamamento já foram apontados como variáveis preditoras significativas em diferentes estudos. Técnicas baseadas em árvores também têm sido eficazes na hierarquização de variáveis em modelos multivariados, facilitando a tomada de decisão em processos seletivos (RINALDI et al., 2022; WASHBURN et al., 2021).

Além disso, a literatura recente tem destacado que essas abordagens oferecem vantagens concretas para a seleção indireta, especialmente em fases iniciais de avaliação genotípica. Isso se deve à possibilidade de incorporar características fisiológicas e morfológicas que, embora secundárias, contribuem significativamente para o desempenho agrônômico. Métodos como o XGBoost, que combinam a estrutura de árvores com técnicas de boosting, têm ampliado ainda mais as aplicações dessas ferramentas, permitindo modelar interações complexas e aumentar a acurácia preditiva em bancos de dados com alta dimensionalidade (CHATTERJEE, S.2000; SIMONOFF, J, 2022).

Em resumo, as árvores de decisão representam uma alternativa robusta e interpretável à análise de trilha, especialmente quando se lida com relações não lineares entre variáveis. Espera-se que a topologia da árvore, em substituição ao diagrama de causa e efeito da análise de trilha, seja igualmente relevante para análise e interpretação da relevância de características auxiliares. Sua aplicabilidade no melhoramento genético é ampla, favorecendo tanto a análise de características auxiliares quanto a predição de traços principais com base em múltiplas fontes de informação. Isso contribui diretamente para o avanço de estratégias mais eficientes e integradas de seleção em programas genéticos modernos.

3. MATERIAIS E MÉTODOS

A fim de avaliar de forma robusta o comportamento dos genótipos em diferentes condições ambientais, foram adotadas análises estatísticas que permitissem isolar e interpretar os efeitos de genótipo, safra e regime hídrico, bem como suas interações.

3.1. Material Experimental

Os dados experimentais envolvendo milho-pipoca neste trabalho são os mesmos utilizados no desenvolvimento da tese do Leite (2023), do programa de pós-graduação em Genética e Melhoramento da Universidade Estadual do Norte Fluminense – UENF. Essas linhagens têm sua origem genética ligada a genótipos desenvolvidos em países da América Latina, adaptados naturalmente a climas tropicais ou temperados (Santos et al., 2021), vide Tabela 1.

Tabela 1 – Descrição de genótipos de milho-pipoca - Leite (2023)

Informações quanto à geração, país de origem, procedência e adaptação climática.

	Genótipos	Pedigree	Origem	Instituição Doadora	Adaptação Climática
T1	L217	IAC 125	Brasil	IAC	Tropical
T2	L381	SAM	EUA	EUA	Temperada
T3	L326	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical
T4	L502	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T5	L292	URUG 298	Uruguai	CIMMYT	Temperada
T6	L503	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T7	L76	Viçosa- Viçosa	Brasil	UFV	Tropical
T8	L291	URUG 298	Uruguai	CIMMYT	Temperada
T9	L273	PARA 172	Paraguai	CIMMYT	Temperada
T10	L328	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical

T11	L61	BRS Ângela	Brasil	EMBRAPA	Tropical
T12	L625	PA 091	Brasil	UEM	Tropical
T13	L688	UENF 14	Brasil	UENF	Tropical
T14	L213	IAC 125	Brasil	IAC	Tropical
T15	L322	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical
T16	L481	SE 013	Brasil	UEM	Tropical
T17	L477	SE 013	Brasil	UEM	Tropical
T18	L684	UENF 14	Brasil	UENF	Tropical
T19	L509	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T20	L501	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T21	L358	PR 023	Brasil	UEM	Tropical
T22	L476	SE 013	Brasil	UEM	Tropical
T23	L472	SE 013	Brasil	UEM	Tropical
T24	L212	IAC 125	Brasil	IAC	Tropical
T25	L204	IAC 125	Brasil	IAC	Tropical
T26	L386	SAM	EUA	EUA	Temperada
T27	L222	IAC 125	Brasil	IAC	Tropical
T28	L691	UENF 14	Brasil	UENF	Tropical
T29	L594	RS 20	Argentina	ISAGRO / AGROSTE	Temperada
T30	L655	ARZM 13 050	Argentina	CIMMYT	Temperada
T31	L391	SAM	EUA	EUA	Temperada
T32	L332	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical

T33	L325	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical
T34	L507	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T35	L263	PARA 172	Paraguai	CIMMYT	Temperada
T36	L219	IAC 125	Brasil	IAC	Tropical
T37	L366	PR 023	Brasil	UEM	Tropical
T38	L394	SAM	EUA	EUA	Temperada
T39	L383	SAM	EUA	EUA	Temperada
T40	L382	SAM	EUA	EUA	Temperada
T41	L330	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical
T42	L652	ARZM 13 050	Argentina	CIMMYT	Temperada
T43	L693	UENF 14	Brasil	UENF	Tropical
T44	L510	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T45	L321	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical
T46	L324	UFV M-2 Barão de Viçosa	Brasil	UFV	Tropical
T47	L513	PA 170 Roxo	Paraguai	CIMMYT	Temperada
T48	L294	URUG 298	Uruguai	CIMMYT	Temperada
T49	L384	SAM	EUA	EUA	Temperada
T50	L220	IAC 125	EUA	IAC	Tropical

USA – Estados Unidos da América; UFV – Universidade Federal de Viçosa; EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária; IAC – Instituto Agrônomo de Campinas; CIMMYT – Centro Internacional de Melhoramento de Milho e Trigo; UEM – Universidade Estadual de Maringá; IPAGRO – Instituto de pesquisas Agrônomicas; e UENF – Universidade Estadual do Norte Fluminense Darcy Ribeiro.-Tabela extraída da tese de leite(2022)

De acordo com Leite (2023): "as características avaliadas foram divididas em três grupos, respectivamente, as agrônômicas, as fisiológicas e as morfológicas do sistema radicular. As características agrônômicas avaliadas, foram: altura média de planta (PH), altura média de espiga (EH), diâmetro médio de colmo (SD), comprimento médio de espiga (EL), diâmetro médio de espiga (ED), número médio de grãos por 17 fileira (GNR), número médio de fileiras por espiga (RNE), número total de grãos por espiga (GNE), massa média de 100 grãos (100GW), rendimento de grãos (GY), capacidade de expansão (PE). As características fisiológicas avaliadas foram: intensidade de cor verde das folhas (estimativa do teor relativo de clorofilas, índice SPAD), teor relativo de antocianinas (ANT) e flavonoides (FLA), índice de balanço de nitrogênio (NBI) e fluorescência da clorofila (Fv/Fm). As características radiculares mensuradas foram o número de raízes de suporte (NBR), o número de raízes da coroa (NCR); a densidade de raízes de suporte (DBR), a densidade das raízes da coroa (DCR), o ângulo de raízes de suporte (ABR) e o ângulo das raízes da coroa (ACR). Foram avaliadas 50 linhagens de milho-pipoca em duas condições hídricas contrastantes: uma bem irrigada (WW) e outra sob estresse hídrico (WS). Os ensaios foram conduzidos nas safras de 2020 e 2021, no Colégio Estadual Agrícola Antônio Sarlo, em Campos dos Goytacazes – RJ, sob coordenadas geográficas 21°34'31" S e 41°54'40" W. Utilizou-se o delineamento em blocos casualizados (DBC), com três repetições, totalizando 150 unidades experimentais. Cada parcela consistiu de uma fileira de 4,40 m, com espaçamento de 0,80 m entre linhas e 0,20 m entre plantas, contendo 23 plantas por parcela. O fornecimento de água foi realizado por irrigação por gotejamento, com um gotejador por planta, permitindo controle preciso da lâmina aplicada. Na condição WW, seguiu-se a recomendação técnica da cultura, enquanto na condição WS a irrigação foi suspensa 15 dias antes da antese masculina, mantendo-se até a colheita. A antese foi determinada com base em dados de experimentos anteriores. A quantidade de água aplicada via irrigação variou entre os tratamentos e anos, havendo redução média de cerca de 47% entre WW e WS. Sensores de umidade do solo (Decagon MPS-6) foram utilizados para monitorar o potencial hídrico, que variou de próximo à capacidade de campo (-0,01 MPa) na condição irrigada até o ponto de murcha permanente (-1,5 MPa) na condição de estresse, em diferentes estádios fenológicos. A área experimental contou ainda com estação meteorológica automática, que registrou temperatura, umidade relativa e radiação fotossinteticamente ativa durante os ensaios". Detalhamentos adicionais sobre os procedimentos metodológicos adotados neste experimento estão disponíveis na tese de Leite (2023), onde o estudo é descrito de forma mais aprofundada.

3.2 Potenciais das Linhagens de Milho-Pipoca e Quantificação da Variabilidade Genética

Com o objetivo de avaliar o potencial genético dos genótipos de milho-pipoca e a quantificação da variabilidade entre elas, foi realizada análise de variância (ANOVA) individual para cada ambiente experimental, considerando os anos de condução e os diferentes regimes hídricos (irrigado e sob estresse). Essa análise foi conduzida com base no delineamento em blocos ao acaso (DBC), apropriado para controlar a variabilidade ambiental e garantir maior precisão na comparação entre os genótipos avaliados.

A análise de variância seguiu o seguinte modelo estatístico:

$$y_{ij} = \mu + G_i + B_j + \varepsilon_{ij}$$

y_{ij} representa o valor observado para a variável em estudo, referente à j-ésima repetição do i-ésimo genótipo;

μ é a média geral;

G_i é o efeito do i-ésimo genótipo;

B_j é o efeito efeito do j-ésimo bloco;

ε_{ij} é o erro experimental associado à observação Y_{ij} , assumido como normalmente e independentemente distribuído com média zero e variância constante ($\varepsilon_{ij} \sim \text{NID}(0, \sigma^2)$).

Com base na ANOVA, verificou-se a existência de variabilidade genética entre os genótipos a partir da hipótese testada foi $H_0: \mu_1 = \mu_2 = \dots = \mu_g$, por meio do teste F, o qual avalia se as diferenças entre as médias dos genótipos são iguais. Adicionalmente, foi realizado o teste de comparação de médias, a saber, o teste de Scott-Knott, a fim de identificar quais genótipos diferem entre si nas características avaliadas. Para facilitar a leitura, os genótipos foram agrupados, com base no teste de Scott-Knott a 5% de significância, sendo representados por letras iguais quando não houve diferença estatística entre eles.

Além disso, foram estimados os componentes de variabilidade genética (ϕ_g) e de variância ambiental (σ^2), a partir dos quadrados médios da ANOVA. Esses valores permitiram o cálculo do coeficiente de determinação genotípica (h^2), que expressa a proporção da variação fenotípica total atribuída às diferenças genéticas entre os genótipos:

$$h^2 = \frac{\hat{\phi}_g}{\frac{\hat{\sigma}^2}{r} + \hat{\phi}_g}$$

onde:

$\hat{\phi}_g$ é a estimativa do componente quadrático que expressa a variabilidade genética entre genótipos;

$\hat{\sigma}^2$ é a estimativa do componente de variância associado ao erro experimental;

r é o número de repetições por genótipo.

Esse coeficiente é fundamental para indicar o grau de confiabilidade na seleção, uma vez que valores mais altos sugerem maior contribuição genética para a variação observada, tornando mais eficiente a identificação de genótipos superiores. Todas as análises estatísticas foram conduzidas com o auxílio do software GENES (CRUZ, 2013).

3.3 Relações entre Características Auxiliares e a Produção e a Capacidade de Expansão de Milho-Pipoca

Para compreender melhor como determinadas características morfoagronômicas se relacionam com os principais atributos de interesse no melhoramento do milho-pipoca, como a produção de grãos e a capacidade de expansão, foram realizadas análises de correlação em diferentes ambientes experimentais. Os dados utilizados foram obtidos nas safras de 2020 e 2021, conduzidas sob duas condições hídricas contrastantes: uma bem irrigada (WW) e outra sob estresse hídrico (WS).

As correlações fenotípicas, genotípicas, ambientais e parciais foram estimadas separadamente para cada ambiente (ano x condição hídrica), permitindo uma avaliação mais detalhada das interações entre as variáveis em contextos distintos. Com base em Cruz (2012), para estimar as correlações supracitadas recomenda-se as análises individuais, e a análise da soma dos caracteres em estudo, de tal forma que as covariâncias, associados a cada fonte de variação possam ser estimados por meio de:

$$Cov(Y_w, Y_{w'}) = \frac{V(Y_w + Y_{w'}) - V(Y_w) - V(Y_{w'})}{2}$$

dessa forma estima-se os coeficientes de correlações por meio de expressões descritas a seguir:

- a) Correlação fenotípica

$$r_f = \frac{PMT_{Y_w Y_{w'}}}{\sqrt{QMT_{Y_w} QMT_{Y_{w'}}}}$$

em que $PMT_{Y_w Y_{w'}}$ é o produto médio associado ao tratamento.

- b) Correlação de ambiente

$$r_f = \frac{PMR_{Y_w Y_w'}}{\sqrt{QMR_{Y_w} QMR_{Y_w'}}$$

em que $PMR_{Y_w Y_w'}$ é o produto médio associado ao resíduo.

c) Correlação genotípica

$$r_g = \frac{\hat{\sigma}_{gxy}}{\sqrt{\hat{\phi}_{gx} \hat{\phi}_{gy}}}$$

em que:

$\hat{\sigma}_{gxy}$ é a estimativa da covariância genotípica entre os caracteres X e Y.

$\hat{\phi}_{gx}$ e $\hat{\phi}_{gy}$ são as estimativas dos componentes quadráticos genotípicos dos caracteres X e Y, respectivamente.

d) Correlação parcial

Numa visão mais simplificada, quando se tem interesse na relação entre duas variáveis, mas desconsiderando a ação de uma terceira variável, pode-se quantificar a correlação por meio de:

$$r_{xy.z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}$$

De maneira generalizada, dado um conjunto de variáveis, faz-se a análise, considerando as matrizes de correlação (R) entre as variáveis especificadas para análise e sua inversa (R^{-1}), conforme definido a seguir:

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1m} \\ & 1 & \dots & r_{2m} \\ & & \dots & \dots \\ & & & 1 \end{bmatrix} \quad R^{-1} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ & c_{22} & \dots & c_{2m} \\ & & \dots & \dots \\ & & & c_{mm} \end{bmatrix}$$

A correlação entre os caracteres i e j, removendo-se a influência dos demais, é assim

$$\text{estimada: } r_{ij.m} = \frac{-c_{ij}}{\sqrt{c_{ii}c_{jj}}}$$

em que m representa o conjunto de caracteres cuja influência na correlação entre i e j foi removida.

O foco esteve voltado para a identificação de relações consistentes entre as características auxiliares — aquelas que não são diretamente selecionadas como objetivo

principal, mas que podem contribuir para a seleção indireta — e os caracteres de maior interesse econômico, como o rendimento e a qualidade dos grãos.

Além de estimar os coeficientes de correlação, foi avaliada a significância das correlações genotípicas por meio do teste t, utilizando os níveis de 5% e 1% de probabilidade.

A hipótese de que o coeficiente de correlação é igual a zero ($H_0: \rho = 0$) foi avaliada pela estatística t, dada por:

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{n - v}$$

em que t está associado a n-v graus de liberdade, sendo v=2 para correlação simples ou v=m+2 para correlação parcial envolvendo duas variáveis e outras m cujos efeitos são removidos.

Esse procedimento foi essencial para garantir a robustez das inferências estatísticas e embasar a interpretação dos resultados com maior segurança.

Todas as análises estatísticas desta etapa foram realizadas utilizando o software GENES (CRUZ, 2013).

3.4 Identificação de Caracteres Auxiliar no Melhoramento da Produção e Capacidade de Expansão do Milho- Pipoca

Este estudo foi conduzido em partes para investigar a relação entre as características agronômicas, fisiológicas e morfológicas em relação com o rendimento de grão e a capacidade de expansão.

3.4.1 Abordagem Fundamentada em Associação Linear: Análise de Trilha

A análise de trilha foi conduzida com base nas correlações fenotípicas obtidas entre os caracteres. Essa técnica permite desdobrar os efeitos diretos e indiretos de variáveis explicativas sobre uma variável resposta, fornecendo uma visão mais detalhada da contribuição de cada característica para o melhoramento.

Para a produção de grãos e para a capacidade de expansão — tratadas separadamente como variáveis principais — foram ajustados modelos distintos, considerando as características morfoagronômicas, fisiológicas e radiculares como variáveis explicativas. As análises foram realizadas individualmente para cada ambiente experimental (ano e condição

hídrica), permitindo a identificação de padrões consistentes ou específicos em diferentes contextos de cultivo.

A matriz de correlações fenotípicas foi inicialmente avaliada quanto à multicolinearidade. Quando identificado alto grau de dependência entre as variáveis explicativas (com base no fator de inflação da variância - VIF), optou-se por estratégias de eliminação sequencial de variáveis redundantes, priorizando a manutenção daquelas com maior relevância biológica e agrônômica. O critério de exclusão considerou a eliminação do caráter com maior VIF, até que os valores ficassem abaixo de um limite aceitável ($VIF < 10$), conforme proposto por Montgomery e Peck (1981).

Para compreender melhor as relações entre as características avaliadas e as variáveis principais de interesse — rendimento de grãos (GY) e capacidade de expansão (PE) —, foi realizada uma análise de trilha com estrutura em cadeias simples. Para a variável GY, foram consideradas separadamente duas análises: a primeira utilizando como variáveis explicativas as características agrônômicas (EL, ED, NGR, GNR, RNE, NGE e GW), e a segunda com as características morfofisiológicas (ANT, CC, FLA, Fv/Fm, SD, EH, PH, ABR, NBR, ACR e NCR). O mesmo procedimento foi adotado para a variável PE, também avaliada com base nas características agrônômicas e, em seguida, nas morfofisiológicas. Essa abordagem permitiu isolar os efeitos diretos e indiretos de cada grupo de características sobre os desfechos de interesse, facilitando a interpretação e identificação de atributos com maior potencial para seleção no melhoramento do milho-pipoca.

A análise parte de um sistema de equações conhecido como equações normais, descrito pela relação:

$$X'X \cdot \hat{\beta} = X'Y$$

Nesse sistema, $X'X$ representa a matriz de correlações entre as 19 variáveis explicativas (uma matriz simétrica), β é o vetor de coeficientes de trilha (efeitos diretos de cada variável sobre GY ou PE), e $X'Y$ é o vetor de correlações entre cada variável explicativa e a variável principal (CRUZ et al., 2012).

A matriz $X'X$ é composta por todas as correlações possíveis entre as variáveis explicativas, enquanto $X'Y$ traz as correlações simples com a variável de interesse (GY ou PE). Por meio desse sistema, é possível estimar os efeitos diretos $\hat{\beta} = (X'X)^{-1}X'Y$ de cada variável, o que contribui para uma compreensão mais aprofundada das influências diretas e indiretas no desempenho das linhagens, especialmente sob diferentes condições hídricas e safras.

A partir da matriz ajustada, foram estimados os efeitos diretos e indiretos por meio da decomposição das correlações genotípicas. Essa abordagem permitiu destacar as variáveis com maior influência direta sobre os caracteres de interesse, apontando caminhos mais seguros para a seleção indireta no melhoramento do milho-pipoca. Com base nas estimativas dos efeitos diretos e indiretos das 17 variáveis explicativas será estimado também o coeficiente de determinação do modelo causal (R^2), é ele quem mede os efeitos das 17 variáveis explicativas sobre GY e PE.

As análises de trilha foram realizadas com o auxílio do software GENES (CRUZ, 2013), com planilhas ajustadas no Microsoft Excel para organização dos dados. O diagrama causal do modelo pode ser observado na Figura 1.

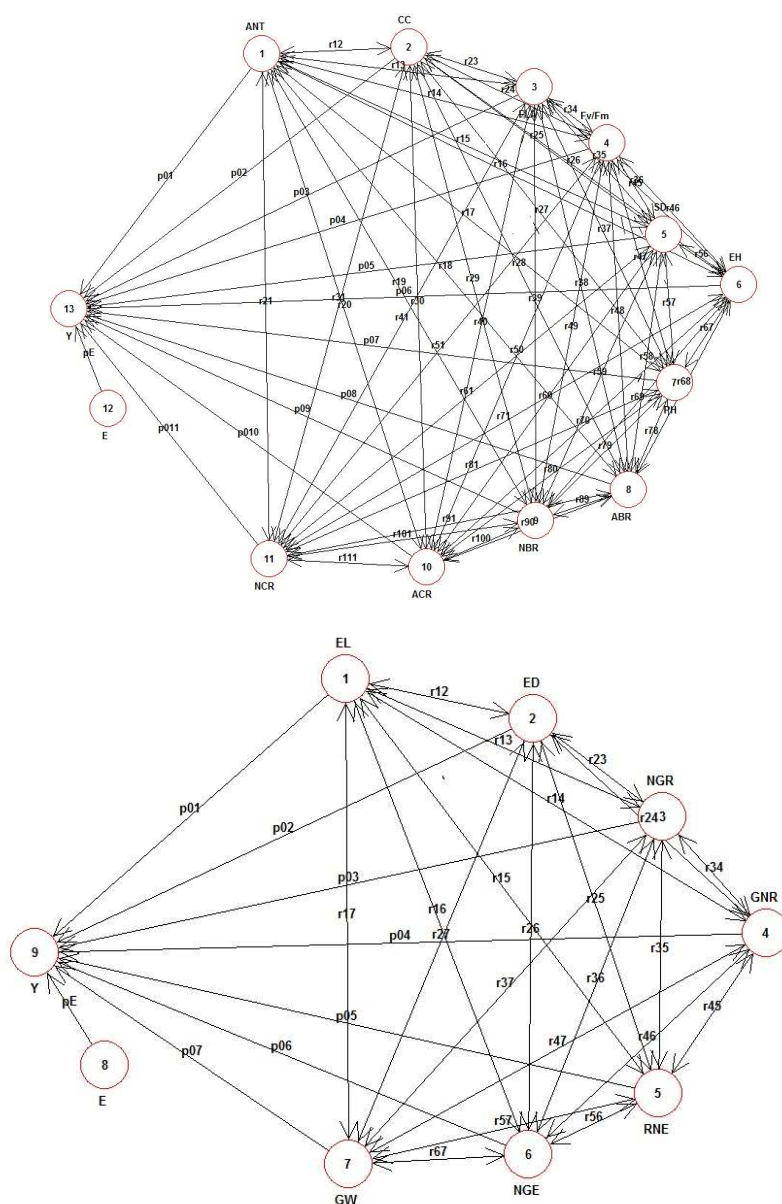


Figura 1 - Diagramas causais dos efeitos das variáveis explicativas sobre a variável dependente Y, representada por GY ou PE. Sendo p_{oi} é o efeito direto da i -ésima variável auxiliar sobre a variável principal e r_{ij} a correlação entre as variáveis explicativas i e j . E representa a variável residual e p_E o efeito direto de E sobre a variável principal Y.

3.4.2 Abordagem Fundamentada em Associação Não-Linear: Árvore de Decisão

Complementando a análise de trilha, este estudo utilizou o método de árvores de decisão como uma abordagem alternativa, não paramétrica, para a identificação de características auxiliares associadas ao rendimento de grãos (GY) e à capacidade de expansão (PE) em milho-pipoca. As árvores foram construídas com base no algoritmo CART (Classification and Regression Trees), metodologia desenvolvida por Breiman et al. (1994), reconhecida por sua habilidade em lidar com relações não lineares e complexas entre variáveis. Neste trabalho foi proposto o seu uso como procedimento alternativo para estudo de associação entre uma variável principal e auxiliares permitindo quantificar tanto o grau de relacionamento quanto a estrutura de árvore, análogo ao diagrama de causa e efeito da análise de trilha, que conduziria a melhores respostas na variável principal representada pela capacidade de expansão e produção do milho pipoca.

Neste estudo, GY e PE foram consideradas como variáveis dependentes em análises separadas. As variáveis explicativas utilizadas como preditoras foram: características agrônomicas (EL, ED, NGR, GNR, RNE, NGE, GW) e características morfológicas e fisiológicas (ANT, CC, FLA, Fv/Fm, SD, EH, PH, ABR, NBR, ACR, NCR). Para cada uma das árvores, foram fixados até 6 nós terminais (ou folhas), a fim de evitar sobreajuste.

O algoritmo CART realiza uma divisão binária recursiva ("recursive binary splitting"), um processo ganancioso que busca a melhor divisão dos dados em cada etapa. Cada divisão visa reduzir o erro de classificação ou predição, agrupando os dados em regiões cada vez mais homogêneas quanto à variável resposta.

O algoritmo usado determina as regiões mutuamente exclusivas e particiona o espaço das variáveis preditoras, fundamentando-se na seguinte estratégia:

- a) Considere uma partição do espaço gerado pelas variáveis preditoras, X, em M regiões, R_1, \dots, R_M

b) Para cada elemento pertencente a R_j , o predictor de Y (que designaremos por \hat{Y}_{R_j}) será a média.

A construção da árvore segue os seguintes passos:

1. Seleciona-se uma variável preditora X_j e um limiar t que divide o espaço das variáveis em duas regiões: $R_1 = \{X : x_j < t\}$ e $R_2 = \{X : x_j \geq t\}$, de forma que o erro de classificação (ou a soma dos quadrados dos resíduos, no caso de regressão) seja minimizado.
2. Entre todos os pares possíveis (j, t) , o algoritmo identifica aquele que gera a menor impureza nos nós resultantes, com base em critérios como a variância residual.
3. O processo é repetido para cada nova região criada, até atingir um critério de parada, que neste estudo foi definido como o limite máximo de 6 nós terminais por árvore.
4. Em cada nó terminal, a predição da variável resposta é a média dos valores observados no subconjunto correspondente de dados.

O desempenho preditivo de cada árvore foi quantificado pelo coeficiente de determinação (R^2), calculado como:

$$R^2 = 1 - \frac{SQR}{SQT}$$

em que SQR é a soma dos quadrados dos resíduos, e SQT é a soma total dos quadrados da variável resposta. Este índice expressa a proporção da variabilidade da variável resposta explicada pelas divisões da árvore.

A utilização de árvores de decisão neste contexto se justifica pela sua capacidade de particionar o espaço das variáveis explicativas de forma hierárquica, revelando interações entre características e identificando aquelas que mais influenciam os caracteres principais (GY e PE). Além disso, a natureza visual e interpretável das árvores favorece a aplicação prática em programas de melhoramento, especialmente na seleção indireta baseada em variáveis morfofisiológicas e agrônomicas.

Todas as análises foram conduzidas no software GENES (CRUZ, 2013), assegurando padronização estatística e confiabilidade dos resultados.

4. RESULTADOS E DISCUSSÃO

4.1. Potenciais de Genótipos de Milho-Pipoca e Variabilidade

Inicialmente foi avaliado o potencial de desempenho do conjunto de genótipos de milho pipoca em relação às variáveis estudadas e a existência de variabilidade para fins de realizar futuras inferências relativas a associação entre características auxiliares e principais no melhoramento genético. Na análise referente à safra de 2020 foi possível identificar diferenças significativas em todas as características avaliadas, considerando nível de 5% de probabilidade (Tabela 2), tanto para as condições de déficit hídrico quanto de irrigação adequada.

Tabela 2- Resumo da análise de variância, das médias estimadas, dos quadrados médios dos tratamentos, bem como das estimativas do coeficiente de variação genético e de determinação genotípico, obtidos a partir da avaliação de características agrônômicas, fisiológicas e morfológicas do sistema radicular de linhagens de milho-pipoca cultivadas sob duas condições hídricas distintas (com e sem estresse hídrico) no ano de 2020.

Características	Condição Hídrica	QMG (GL=29)	\bar{X}	CV(%)	h^2
SD (mm)	ww	156,38**	11,31	2,35	98,58
	ws	1076,54**	128,96	2,67	98,89
EH (cm)	ww	780,62**	90,89	5,62	96,66
	ws	660,40**	80,84	5,00	97,52
PH (cm)	ww	3,67**	12,28	7,71	75,58
	ws	4,73**	12,89	7,61	79,70
EL(cm)	ww	7,40**	10,95	6,58	92,97
	ws	10,78**	9,70	7,40	95,21
ED (mm)	ww	15,43**	27,79	5,43	85,23
	ws	23,16**	25,18	5,73	91,01
GNE (unid.)	ww	48,02**	13,11	11,61	87,51
	ws	34,82**	11,70	9,42	94,92
RNE (unid.)	ww	9687,24**	299,63	6,21	89,39
	ws	8470,55**	220,07	5,51	95,56
GNR (unid.)	ww	6,25**	22,89	8,16	92,72
	ws	9,37**	18,76	8,15	93,27
100GW (g)	ww	11,54**	10,90	6,52	95,62
	ws	8,51**	9,18	6,00	96,43
GY(kg. ha ⁻¹)	ww	1118030,15**	1519,63	6,23	96,73
	ws	195448,29**	461,26	15,05	97,53
PE (g. ml ⁻¹)	ww	69,53**	22,19	6,23	97,25
	ws	66,84**	17,30	7,35	97,57
CC	ww	145,62**	29,75	7,17	96,87

FLA	ws	163,27**	14,98	9,64	98,72
	ww	0,0028**	0,23	8,02	88,26
ANT	ws	0,0031**	0,26	7,28	88,09
	ww	0,056**	1,22	3,89	95,98
Fv/Fm	ws	0,056**	1,22	2,69	98,07
	ww	0,004**	0,72	3,36	83,51
ABR (°)	ws	0,005**	0,70	4,82	77,52
	ww	113,87**	33,52	10,83	88,42
ACR(°)	ws	181,68**	35,80	14,55	85,07
	ww	256,16**	41,35	8,67	94,98
NBR (unid.)	ws	205,86**	42,07	8,36	93,99
	ww	26,22**	11,83	9,46	95,23
NCR (unid.)	ws	25,50**	10,79	18,35	84,61
	ww	119,91**	16,82	14,11	95,30
	ws	64,41**	16,05	14,46	91,63

WS – Condição com déficit hídrico; WW – Irrigação adequada; CVe – Coeficiente de variação genético; EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. ** e * indica diferença significativa em nível de 1% e 5% pelo teste F, respectivamente; ns: ausência de diferença significativa pelo teste F em 5% de probabilidade.

Também, na análise referente à safra de 2021, foi possível identificar diferenças significativas em todas as características avaliadas, considerando um nível de 5% de probabilidade (Tabela 3).

Tabela 3 Resumo da análise de variância, das médias estimadas, dos quadrados médios dos tratamentos, bem como das estimativas do coeficiente de variação genético e de determinação genotípico, obtidos a partir da avaliação de características agrônomicas, fisiológicas e morfológicas do sistema radicular de linhagens de milho-pipoca cultivadas sob duas condições hídricas distintas (com e sem estresse hídrico) no ano de 2021.

Características	Condição Hídrica	QMG (GL=29)	\bar{X}	CV(%)	h ²
SD (mm)	ww	9,42**	14,36	5,23	94,01
	ws	9,68**	12,79	5,30	95,25
EH (cm)	ww	485,28**	93,83	3,50	97,78
	ws	583,17**	90,30	3,09	98,66
PH (cm)	ww	1116,81**	160,06	2,99	97,94
	ws	1348,96**	154,99	2,99	98,40
EL (cm)	ww	400,80**	12,62	5,23	94,66
	ws	10,57**	12,07	5,84	95,28
ED (mm)	ww	29,42**	27,84	5,96	90,64
	ws	25,88**	26,06	7,28	86,09
GNE (unid.)	ww	43,47**	22,78	10,46	86,92
	ws	89,63**	22,53	11,39	92,65
RNE (unid.)	ww	12742,47**	288,97	14,20	86,78

GNR (unid.)	ws	24329,87**	270,76	16,62	91,67
	ww	6,34**	12,60	8,35	82,55
100GW (g)	ws	8,71**	11,78	8,99	87,12
	ww	11,37**	9,99	12,99	85,17
GY (kg.ha ⁻¹)	ws	7,99**	9,14	12,60	83,36
	ww	803262,01**	1285,04	13,45	96,28
PE (g.ml ⁻¹)	ws	439000,45**	812,07	17,19	95,56
	ww	58,62**	22,15	7,72	95,00
CC	ws	48,61**	18,75	5,79	97,57
	ww	150,75**	32,44	6,70	96,86
FLA	ws	99,33**	14,16	8,96	98,37
	ww	0,08**	0,95	7,42	94,18
ANT	ws	0,09**	1,09	4,92	96,89
	ww	0,0029**	0,063	26,44	90,37
Fv/Fm	ws	0,004**	0,15	10,69	93,54
	ww	0,005**	0,67	7,63	55,53
ABR (°)	ws	0,009**	0,66	8,07	68,40
	ww	251,42**	46,37	9,45	92,36
ACR(°)	ws	251,23**	42,51	14,72	84,40
	ww	208,78**	47,40	8,79	91,68
NBR (unid.)	ws	302,03**	44,6	10,18	93,17
	ww	33,96**	12,05	11,88	93,96
DCR (unid.)	ws	15,83**	10,75	17,96	76,41
	ww	141,76**	22,58	14,62	92,30
	ws	67,09**	20,72	10,63	92,75

WS – Condição com déficit hídrico; WW – Irrigação adequada; CVe – Coeficiente de variação genético; EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. ** e * indica diferença significativa em nível de 1% e 5% pelo teste F, respectivamente; ns: ausência de diferença significativa pelo teste F em 5% de probabilidade.

A presença de variabilidade genética entre os genótipos de milho avaliados mostra que existe bom potencial para identificar e selecionar linhagens superiores, mesmo sob diferentes condições hídricas. Isso é um indicativo importante para o melhoramento, já que a diversidade entre os genótipos é o ponto de partida para qualquer processo seletivo eficiente.

Observa-se que, em ambas as safras e ambientes (bem irrigado e sob estresse hídrico), várias características apresentaram valores elevados de coeficiente de determinação genotípico. Esse resultado é bastante animador, pois indica que essas características são fortemente influenciadas pelos fatores genéticos, o que aumenta as chances de sucesso na

seleção de genótipos ou de indivíduos em populações advindas do inter cruzamento das linhagens estudadas. Como apontam Mastrangelo et al. (2012) e Li et al. (2024), quando uma característica apresenta alto coeficiente de determinação genotípico, ela tende a ser mais estável e previsível, tornando-se um alvo confiável para programas de seleção, inclusive em condições adversas, como a seca.

Em alguns casos, foi possível notar que o coeficiente de determinação genotípico de certas características aumentou quando as plantas estavam sob estresse hídrico, o que pode indicar que essas variáveis se tornam ainda mais úteis na seleção de genótipos tolerantes à seca. Em condições sob estresse hídrico as diferenças relativas aos atributos que determinam o comportamento de tolerância à falta de água são realçadas permitindo a escolha dos genótipos, mas apropriados sob esta condição.

De maneira mais simples pode-se pensar em algumas possíveis estratégias para fins de obter materiais tolerantes à seca. Uma delas é conduzir um único programa em condições favoráveis e, no final do ciclo de melhoramento, recorrer a ensaios em condições de estresse para constatação de que genótipos também adaptados à esta condição estejam presentes na população melhorada. Outra estratégia é a condução de programa de melhoramento específico para estas condições de estresse onde a seleção, em cada fase do programa de melhoramento, seja efetuada com a observância do desempenho do conjunto de genótipos nestas condições. Os resultados encontrados apontam que o estabelecimento de programas de melhoramento sob esta condição específica é uma estratégia adequada e deve ser recomendada quando o objetivo for a obtenção de cultivares tolerantes.

Outro ponto que merece destaque é a qualidade do experimento, avaliada por meio dos coeficientes de variação (CVe). Em ambas as safras, os valores observados foram baixos, o que reforça a precisão dos dados obtidos e a confiabilidade das diferenças observadas entre os genótipos. Esse tipo de resultado garante segurança na interpretação e reforça a qualidade da condução experimental.

Na sequência, são apresentadas as Tabelas com os resultados de capacidade de expansão e rendimento de grãos que permite particularizar se um genótipo apresenta, ou não, bom desempenho em condições adequadas e desfavoráveis

Tabela 4. Agrupamento de genótipos (T1 a T50) pelo teste de Scott-Knott ($p \geq 0,05$) para rendimento de grãos (kg ha^{-1})

WW no ano de 2020		
Grupo de Médias	Genótipos	Médias kg ha^{-1}
a	T1, T12, T33, T45	2119,11 – 2403,36
b	T5, T21, T26, T31, T37, T47	1877,17 – 1951,70
c	T9, T28, T32, T35, T38, T39, T42, T48, T49	1416,29 – 1777,36
d	T4, T6, T10, T11, T13, T15, T18, T24, T25, T29, T30, T41, T43, T44, T50	1045,45 – 1369,32
e	T2, T8, T14, T17, T20, T36, T40, T46	782,33 – 950,38
f	T3, T7, T16, T19, T22, T23, T27, T34	259,73 – 691,24

WS no ano de 2020		
Grupo	Genótipos	Médias (kg ha^{-1})
a	T11, T17, T34	1084,46 – 1201,21
b	T32	920,71
c	T12, T24, T38	738,64 – 785,67
d	T13, T14, T22, T26, T27, T33, T36, T39, T44	550,73 – 672,35
e	T1, T9, T19, T23, T37, T47, T48	426,14 – 501,49
f	T3, T5, T16, T20, T21, T29, T30, T35, T49, T50	337,12 – 401,69
g	T2, T4, T6, T7, T8, T10, T15, T18, T25, T28, T31, T40, T41, T42, T43, T45, T46	160,15 – 309,62

Genótipos agrupados pela mesma letra não diferem em média, pelo teste de Scott-Knott. A ordem de tratamentos são os mesmos apresentados na Tabela 1.

Ao comparar o agrupamento dos tratamentos pelo teste de Scott-Knott ($p \geq 0,05$) para o rendimento de grãos (GY) em 2020, tanto em condição bem irrigada (WW) quanto sob estresse hídrico (WS), percebe-se que nenhum genótipo conseguiu se manter entre os melhores nas duas situações.

Na condição bem irrigada, destacaram-se os genótipos T1, T45, T12 e T33, que formaram o grupo **a**, além de T26, T47, T21, T31, T5 e T37, que compuseram o grupo **b**. No entanto, quando avaliados sob estresse hídrico, esses mesmos genótipos migraram para grupos intermediários ou inferiores (**d**, **e**, **f** e **g**), indicando uma queda de desempenho considerável.

Por outro lado, sob estresse hídrico, quem se destacaram foram os genótipos T34, T17 e T11, reunidos no grupo **a**, além de T32, classificado no grupo **b** com média de 920,71 kg ha⁻¹. Curiosamente, esses genótipos não haviam figurado entre os melhores quando cultivados em condição bem irrigada.

Esses resultados mostram de forma clara como a resposta dos genótipos varia conforme o ambiente. Em outras palavras, uma linhagem que se sobressai em um cenário mais favorável pode não repetir o mesmo desempenho quando submetido ao estresse, e vice-versa.

Tabela 5. Agrupamento de genótipos (T1 a T50) pelo teste de Scott-Knott ($p \geq 0,05$) para capacidade de expansão (mL g⁻¹)

WW no ano de 2020		
Grupo de Médias	Genótipos	Médias
a	T5, T6, T8, T11, T41, T43	29,08 – 31,67
b	T1, T14, T17, T27, T32, T38, T40, T50	25,16 – 27,33
c	T2, T7, T16, T20, T21, T23, T24, T29, T31, T36, T47	21,83 – 23,50
d	T3, T4, T9, T10, T12, T18, T19, T22, T30, T33, T35, T37, T42, T44, T45, T46	19,42 – 21,00
e	T13, T25, T26, T28, T34, T39, T48, T49	15,17 – 18,17
f	T15	12,33

WS no ano de 2020		
Grupo	Genótipos	Médias (mL g⁻¹)
a	T7	27,85
b	T4, T6, T8, T16, T18, T20, T22, T26, T32, T38, T40, T46, T47, T49	21,55 – 24,27
c	T9, T11, T12, T13, T19, T21, T27, T33, T45	17,86 – 20,00
d	T5, T14, T25, T29, T31, T36	15,28 – 17,26
e	T2, T3, T10, T15, T17, T24, T28, T30, T34, T35, T37, T39, T41, T43, T44	12,14 – 14,80
f	T1, T23, T42, T50	10,76 – 11,07
g	T48	3,57

Genótipos agrupados pela mesma letra não diferem em média, pelo teste de Scott-Knott. A ordem de tratamentos são os mesmos apresentados na Tabela 1.

A capacidade de expansão dos grãos mostrou-se bastante sensível às condições hídricas avaliadas, refletindo o impacto direto do ambiente no desempenho dos genótipos. Pela análise de agrupamento obtida com o teste de Scott-Knott ($p \geq 0,05$) observou-se que, na condição bem irrigada (WW), seis linhagens se destacaram no grupo de maior desempenho (**a**), T6, T5, T41, T43, T11 e T8. Logo abaixo, o grupo **b** reuniu genótipos, ainda considerados de bom desempenho, enquanto os demais grupos apresentaram valores progressivamente menores, chegando a 12,33 mL g⁻¹ para o genótipo T15 no grupo **f**.

Sob estresse hídrico (WS), o padrão de agrupamento foi bastante diferente. Apenas o genótipo T7 permaneceu isolado no grupo **a**, atingindo 27,85 mL g⁻¹. O grupo **b** incluiu um conjunto mais amplo de materiais, enquanto os grupos seguintes (**c** a **g**) apresentaram reduções graduais, culminando em 3,57 mL g⁻¹ para o genótipo T48 no grupo **g**.

Ao comparar os grupos de melhor desempenho nas duas condições, fica evidente que nenhum genótipo se manteve simultaneamente no grupo **a** em WW e WS. Ainda assim, destacam-se os genótipos T6 e T8, que apresentaram desempenho consistente: figuraram no grupo **a** em condição bem irrigada e permaneceram no grupo **b** sob estresse hídrico, demonstrando maior estabilidade e potencial de uso em ambientes contrastantes.

Tabela 6. Agrupamento de genótipos (T1 a T50) pelo teste de Scott-Knott ($p \geq 0,05$) para rendimento de grãos (kg ha⁻¹)

WW no ano de 2021		
Grupo de Médias	Genótipos	Médias
a	T1, T12, T33, T45	2119,11 – 2403,36
b	T5, T21, T26, T31, T37, T47	1877,17 – 1951,70
c	T9, T28, T32, T35, T38, T39, T42, T48, T49	1416,29 – 1777,36
d	T4, T6, T10, T11, T13, T15, T18, T24, T25, T29, T30, T41, T43, T44, T50	1045,45 – 1369,32
e	T2, T8, T14, T17, T20, T36, T40, T46	782,33 – 950,38
f	T3, T7, T16, T19, T22, T23, T27, T34	259,73 – 691,24

WS no ano de 2021		
Grupo	Tratamentos	Média
a	T2, T4, T39	1587,14 – 1794,70
b	T6, T35, T37, T44	1288,15 – 1414,24

c	T11, T17, T28, T31, T47, T49, T50	1026,79 – 1196,24
d	T9, T10, T14, T15, T21, T24, T30, T32, T33, T34, T36, T38, T41, T46	791,67 – 977,32
e	T12, T13, T18, T19, T22, T23, T25, T26, T27, T40, T42, T43	484,97 – 691,43
f	T1, T3, T5, T7, T8, T16, T20, T29, T45, T48	85,23 – 432,34

Genótipos agrupados pela mesma letra não diferem em média, pelo teste de Scott-Knott. A ordem de tratamentos são os mesmos apresentados na Tabela 1.

Comparando apenas os grupos **a** e **b** do agrupamento estabelecido pelo teste Scott-Knott verifica-se que apenas o genótipo T37 foi comum se destacando com bom desempenho em ambas as condições de fornecimento de água. Resultado de maior impacto é a comparação dos genótipos do grupo **a** WW 2021, de melhor desempenho com o grupo **f**, WS 2021, em que os tratamentos T1 e T45 se descaram indicando que aqueles de melhor desempenho em uma condição de destacaram como de pior desempenho em outra condição.

A situação é mais amena quando o reconhecimento do desempenho é feito em ensaios com estresse em relação a ensaios com condições mais favoráveis. Assim, os genótipos de bom desempenho em condições de estresse hídrico (T4, T2, T39 que constituem o grupo **a** em WS no ano de 2021) tiveram desempenho mediano em condições mais adequadas uma vez que T4 ocupou o grupo **d**, T2 o grupo **e**, sendo o mais prejudicado dos três, e T39 o grupo **c**, ou seja em WW no ano 2021.

Tabela 7. Agrupamento de genótipos (T1 a T50) pelo teste de Scott-Knott ($p \geq 0,05$) para capacidade de expansão (mL g^{-1})

WW no ano de 2021		
Grupo de Médias	Genótipos	Médias
a	T5, T6, T8, T11, T41, T43	29,08 – 31,67
b	T1, T14, T17, T27, T32, T38, T40, T50	25,17 – 27,33
c	T2, T7, T16, T20, T21, T23, T24, T29, T31, T36, T47	21,83 – 23,50
d	T3, T4, T9, T10, T12, T18, T19, T22, T30, T33, T35, T37, T42, T44, T45, T46	19,42 – 21,00
e	T13, T25, T26, T28, T34, T39, T48, T49	15,17 – 18,17
f	T15	12,33

WS no ano de 2021		
Grupo	Tratamentos	Média
a	T37	30,33
b	T24, T43	26,50 – 28,00
c	T3, T10, T12, T14, T20, T28, T44, T47, T48, T50	21,75 – 24,00
d	T27, T32, T34, T39, T40, T46	19,83 – 21,33
e	T2, T4, T6, T7, T8, T16, T26, T29, T30, T36	17,83 – 19,33
f	T19, T21, T33, T38, T41, T42	16,38 – 17,44
g	T1, T5, T9, T15, T17, T18, T22, T25, T31, T45, T49	14,33 – 16,00
h	T11, T13, T23, T35	11,46 – 13,25

Genótipos agrupados pela mesma letra não diferem em média, pelo teste de Scott-Knott. A ordem de tratamentos são os mesmos apresentados na Tabela 1.

O teste de Scott- Knott a 5% de significância permitiu a formação de vários grupos distintos para as variáveis rendimento de grão e capacidade de expansão. Algumas linhagens compuseram grupos superiores, com médias significativamente maiores. O que nos permitem destacá-los como promissores para o cultivo em ambientes com disponibilidade hídrica limitada.

Novamente destacamos um tipo de interação complexa das linhagens em relação às condições de fornecimento de água. Assim, pode-se verificar o comportamento do T11 pertencente ao grupo **a**, de bom desempenho, nas condições WW no ano de 2021 passando a ocupar a posição do último grupo, de pior desempenho, na condição WS no ano de 2021. Apesar desta situação alarmante quanto à seleção de linhagens superiores para diferentes condições ambientais, também há resultado que demonstram sucesso quando a escolha é feita sobre linhagens mais versáteis. Este é o caso dos genótipos T43 e T24 que ocuparam os grupos **a** e **b**, WS no ano de 2021, e os grupos **a** e **c** em condição WW no ano 2021.

Os resultados apontaram uma expressiva interação genótipo \times ambiente, reforçando que a resposta dos genótipos avaliados é fortemente dependente do ambiente e da condição de cultivo.

Para o rendimento de grãos, verificou-se que, em 2020, os genótipos T1, T45, T12 e T33 ocuparam o grupo de maior desempenho em condição bem irrigada (WW), atingindo médias entre 2.119,11 e 2.403,36 kg ha⁻¹, seguidos por T26, T47, T21, T31, T5 e T37, também com valores elevados. No entanto, sob estresse hídrico (WS), esses mesmos materiais

migraram para grupos intermediários ou inferiores (d, e, f e g), enquanto genótipos como T34, T17 e T11 se destacaram, apresentando médias entre 1.201,21 e 1.084,46 kg ha⁻¹. Esse padrão se repetiu em 2021: materiais de alto rendimento sob irrigação, como T1 e T45, passaram a compor grupos de pior desempenho sob estresse hídrico, enquanto genótipos como T4, T2 e T39 lideraram o rendimento em WS, mas ocuparam grupos medianos em WW. Apenas T37 apresentou maior estabilidade relativa, permanecendo em grupos superiores em ambas as condições em 2021.

Esses resultados estão de acordo com estudos que mostram a sensibilidade do rendimento de grãos à restrição hídrica, com reduções que podem alcançar 30,61% (SANTOS et al., 2021) e até 51% em condições de seca severa (KAMPHORST et al., 2020). As variações observadas entre as linhagens são reflexo não apenas do efeito ambiental, mas também de diferenças genéticas intrínsecas, uma vez que a matriz de variância e covariância genética pode mudar entre ambientes, modificando as correlações genéticas entre caracteres como rendimento e teor de amido (HUDSON et al., 2021; LOPEZ-CRUZ et al., 2023).

Quanto à capacidade de expansão, o impacto ambiental também foi evidente. Em WW 2020 e WW 2021, os genótipos T6, T5, T41, T43, T11 e T8 integraram o grupo superior. Esses valores situam-se na faixa considerada como expansão média, pois, de acordo com Machado (2021), valores em torno de 30 mL g⁻¹ indicam um nível moderado de expansão, aceitável para determinados mercados, mas ainda aquém do padrão de excelência comercial, que corresponde a valores superiores a 40 mL g⁻¹ (MACHADO, 2021; CARANHATO et al., 2022). Sob estresse hídrico em 2020, o agrupamento mudou: apenas T7 permaneceu isolado no grupo de maior desempenho (27,85 mL g⁻¹), enquanto T6 e T8 foram alocados no grupo b, mostrando estabilidade relativa. Em WS 2021, T37 foi o único no grupo a (30,33 mL g⁻¹), enquanto T43 e T24 ocuparam os grupos a e b, mantendo bom desempenho, mesmo sob déficit hídrico.

Estudos apontam que o estresse hídrico pode reduzir a capacidade de expansão em aproximadamente 3,5% (SANTOS et al., 2021) e, em casos mais severos, até 55% (KAMPHORST et al., 2020). Essa redução está relacionada a alterações fisiológicas, como o esverdeamento das folhas, que afetam a qualidade do grão e a capacidade de estourar (KAMPHORST et al., 2020).

Além disso, diferenças genéticas explicam a resiliência de alguns genótipos, permitindo que mantenham taxas de expansão mais elevadas mesmo sob limitação de água (OLIVEIRA et al., 2024). A interação genótipo × ambiente tem peso semelhante ao efeito

genético para características agronômicas importantes, incluindo a capacidade de expansão (HUDSON et al., 2021), o que justifica a variação significativa observada entre os tratamentos.

De forma geral, as análises revelam que genótipos de maior desempenho sob irrigação não necessariamente mantêm o patamar em condições de estresse hídrico. Em contrapartida, alguns genótipos, como T6, T8, T43 e T24, demonstraram maior consistência, mantendo boa capacidade de expansão em ambientes contrastantes, o que os tornam promissores para programas de melhoramento voltados a regiões com disponibilidade hídrica variável. Esses achados reforçam que, para se obter cultivares de milho-pipoca com alto rendimento e qualidade tecnológica estável, é indispensável avaliar os materiais em diferentes condições ambientais, valorizando tanto o potencial produtivo quanto a resiliência genética.

O comportamento diferencial dos genótipos de milho em ambientes com e sem déficit hídrico foi avaliado por meio de análises conjuntas, procedimento já explorado por Leite (2022) em sua tese. A significância da interação genótipo \times ambiente ($G \times WC$) evidencia que os genótipos respondem de forma distinta nos diferentes ambientes avaliados. Esse comportamento foi constatado para os caracteres estudados nas distintas condições hídricas e nas diferentes safras.

4.2 Associações Entre Caracteres Agronômicos Auxiliares e a Capacidade de Expansão e Rendimento de Grãos

Nesta seção são apresentados os resultados das correlações fenotípicas simples e parciais obtidas para as safras de 2020 e 2021, avaliadas sob duas condições hídricas: bem irrigada (WW) e sob déficit hídrico (WS). As demais correlações, não discutidas neste capítulo, encontram-se disponíveis no APÊNDICE A.

Tabela 8. Estimativas de correlações fenotípicas simples e parciais entre as características auxiliares e o rendimento de grãos (GY) e capacidade de expansão (PE) nas condições WW e WS, na safra de 2020

Caráter	WW				WS			
	GY		PE		GY		PE	
	r	r parc	r	r parc	r	r parc	r	r parc
SD	0,38**	0,33*	-0,08	0,1	0,27*	0,14	0,09	0,05
EH	0,2*	0,15	-0,21**	-0,15	0,07	-0,07	-0,22**	-0,04
PH	0,02	-0,11	-0,16*	0,11	-0,02	0,05	-0,2*	-0,05
CC	0,13	-0,03	-0,06	-0,09	0,33*	0,29*	-0,14	-0,15
ANT	-0,2*	-0,03	0,23**	0,35**	-0,15	-0,08	-0,12	-0,1
FLA	0,21*	0,24**	0,06	0,13	0,07	0,09	-0,3**	-0,22**
Fv/Fm	0,01	0,03	0,0	0,02	0,03	0,29	0,29**	0,20*
ABR	-0,04	-0,02	-0,01	-0,04	0,01	-0,04	-0,0	0,01
NBR	-	-0,12	-0,28**	-0,37**	0,14	0,25*	-0,1	-0,1
ACR	0,16*	0,05	-0,01	-0,03	0,18*	0,18*	0,19*	0,20*
NCR	0,13	0,01	-0,03	-0,05	0,05	-0,04	-0,27**	-0,12
EL	0,34*	0,10	-0,0	-0,02	0,26*	0,16	0,07	-0,01
ED	0,23*	0,15	-0,37**	-0,27**	0,26*	0,08	0,04	-0,04
GNR	0,41*	0,8	0,05	-0,02	0,27*	-0,16*	0,09	0,01
RNE	0,02	0,04	-0,15	-0,01	0,22*	-0,16	0,01	-0,03
GNE	0,38*	-0,04	-0,03	0,03	0,35*	0,19*	0,08	0,03
100GW	0,25*	0,12	-0,21**	-0,0	0,14	0,1	0,13	0,14

EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da

coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Uma interpretação mais geral é que associações que podem ser significativas em uma determinada condição podem não ser em outra condição. Esta interpretação é válida tanto para GY quanto para PE. Isto indica que as condições ambientais diferenciadas acionam diferentes mecanismos fisiológicos ou estruturas morfológicas de forma que a importância de certas características é alterada para suprir as demandas das condições específicas. Assim, este resultado já indica que o melhoramento não deve generalizar a importância de características auxiliares tomando por base ensaios de uma condição ambiental diferente daquela em que se pretende praticar a seleção.

De maneira geral as correlações parciais não indicaram mudanças de magnitude e sinal das associações simples obtidas. Assim, para este conjunto de variáveis, o impacto de outras variáveis sobre a associação de interesse foi pequeno. Um único resultado mais discrepante foi entre GNR e GY, em WS 2020, que cuja correlação parcial indicou significância, mas com sinal contrário ao da correlação simples. Assim, ao utilizar esta característica auxiliar, recomenda-se estudos adicionais, fundamentados em análise de trilha e similares, para realmente quantificar o efeito direto e indireto desta associação.

Em relação às características agrônômicas (EL, ED, NGR, GNR, RNE, NGE, GW) e características morfológicas e fisiológicas (ANT, CC, FLA, Fv/Fm, SD, EH, PH, ABR, NBR, ACR, NCR), observou-se que o grupo de características agrônômicas apresentou maior importância e consistência na associação com o rendimento de grãos (GY), quando comparado ao grupo de características morfofisiológicas.

De modo geral, as variáveis agrônômicas exibiram correlações fenotípicas mais elevadas e significativas em ambas as condições hídricas, bem irrigada (WW) e sob estresse hídrico (WS). Entre elas, destacaram-se o número de grãos por fileira (GNR), o número de grãos por espiga (GNE), o comprimento de espiga (EL) e o diâmetro de espiga (ED), todas com associações positivas e significativas com o rendimento de grãos. Contudo, o diâmetro de colmo (SD) merece destaque especial, pois apresentou correlações positivas e significativas com GY em WW ($r=0,38^{**}$) e manteve associação consistente em WS ($r=0,27^{**}$). Esse resultado indica que plantas com colmos mais espessos tendem a sustentar melhor a formação e enchimento das espigas, resultando em maiores rendimentos, independentemente do nível de disponibilidade hídrica.

Por outro lado, as variáveis morfofisiológicas apresentaram influência mais modesta e pontual, com algumas associações significativas apenas em determinados contextos, sem se manterem consistentes entre os ambientes avaliados.

Assim, pode-se concluir que o grupo de caracteres agronômicos foi o mais importante para explicar o rendimento de grãos, e, dentro dele, o diâmetro de colmo (SD) se destaca como um caráter auxiliar promissor a ser considerado na seleção de genótipos com maior potencial produtivo, mesmo sob condições de estresse hídrico.

Com base na matriz de correlações fenotípicas na safra 2020 condição WW (Tabela 8), foi possível identificar associações significativas entre diversas características morfológicas, fisiológicas e agronômicas avaliadas. Considerando o nível de 5% e 1% de probabilidade pelo teste t, observou-se que o diâmetro do colmo (SD) apresentou correlação (simples e parcial) positiva significativa com o rendimento de grãos ($r = 0,3765^{**}$) e ($r_{\text{parc}} = 0,33^{**}$), sugerindo que essa característica pode estar associada ao desempenho produtivo das linhagens avaliadas. O teor relativo de flavonoides (FLA) também se destacou, apresentando correlação positiva e significativa com o rendimento ($r = 0,21^{**}$ e $r_{\text{parc}} = 0,24^{**}$). Esse resultado sugere que genótipos com maior concentração de flavonoides podem contribuir de forma direta para o aumento da produção de grãos.

Quando se analisa a capacidade de expansão (PE), ainda em WW, percebe-se um efeito positivo e significativo do teor relativo de antocianinas (ANT), com $r = 0,22^*$ e $r_{\text{parc}} = 0,35^*$. Ou seja, genótipos com maiores teores de antocianinas tendem a apresentar grãos com melhor capacidade de expansão. Em contrapartida, o número de raízes de suporte (NBR) mostrou uma correlação negativa significativa com a expansão ($r = -0,28^*$ e $r_{\text{parc}} = -0,37^*$), indicando que plantas com maior número de raízes de suporte tendem a apresentar menor capacidade de expansão. O diâmetro de espigas (ED) também se destacou por apresentar uma relação negativa e significativa com a capacidade de expansão (PE). Os valores observados foram de $r_{\text{simples}} = -0,37^{**}$ e $r_{\text{parcial}} = -0,27^{**}$. Isso significa que, nesse ambiente favorável, espigas com maior diâmetro tendem a estar associadas a uma menor capacidade de expansão dos grãos, possivelmente refletindo uma relação de compensação entre características de produtividade e qualidade de pipoca.

Para a mesma safra, sob condição de estresse hídrico, começando pelo rendimento de grãos, observa-se que o teor relativo de clorofilas (CC) apresentou correlações positivas e significativas com GY, com $r = 0,33^{**}$ e $r_{\text{parc}} = 0,29^{**}$. Isso indica que plantas capazes de manter níveis mais altos de clorofila em condições de estresse tendem a ter maior produtividade.

O número de raízes de suporte (NBR) também se mostrou relevante, apresentando correlações significativas ($r = 0,14^{**}$ e $r \text{ parc} = 0,25^{**}$) com o rendimento. Esse resultado sugere que sistemas radiculares mais resistentes e adaptativos ajudam a planta a explorar melhor o solo em busca de água e nutrientes, refletindo positivamente na produção.

O ângulo de raízes da coroa (ACR) destacou-se de forma semelhante, com correlações de $r = 0,18^*$ e $r \text{ parc} = 0,818$, ambas significativas, reforçando a importância da arquitetura radicular como um mecanismo adaptativo sob déficit hídrico.

Entre os componentes diretos de produção, o número de grãos por fileira (GNR) mostrou correlação simples positiva com GY ($r = 0,27^*$), mas apresentou correlação parcial negativa ($r \text{ parc} = -0,16^*$). Isso indica que, isoladamente, o aumento do número de grãos por fileira tende a ser vantajoso, mas, quando controlados os efeitos de outras variáveis, pode estar associado a ajustes que comprometem o rendimento.

O número de grãos por espiga (GNE), por sua vez, manteve uma relação positiva com GY ($r = 0,35^*$ e $r \text{ parc} = 0,19^*$), evidenciando sua importância direta para a produtividade. Já a fluorescência de clorofilas (Fv/Fm) apresentou uma correlação simples menor ($r = 0,03^*$), mas com um efeito parcial mais expressivo ($r \text{ parc} = 0,29^*$), indicando que plantas com maior eficiência no uso da luz tendem a manter melhores níveis de rendimento mesmo em condições de déficit.

Quando observamos a capacidade de expansão (PE), o cenário é um pouco diferente. O teor relativo de flavonoides (FLA) apresentou correlações negativas e significativas com PE ($r = -0,30^{**}$ e $r \text{ parc} = -0,22^{**}$), sugerindo que plantas que investem mais na produção dessa característica sob estresse tendem a apresentar grãos com menor capacidade de expansão.

Por outro lado, a fluorescência de clorofilas (Fv/Fm) mostrou-se positivamente associada à expansão ($r = 0,29^*$ e $r \text{ parc} = 0,20^*$), indicando que plantas mais eficientes no aproveitamento da luz mantêm uma melhor qualidade do grão, mesmo sob déficit hídrico.

O ângulo de raízes da coroa (ACR) também se relacionou com a capacidade de expansão, apresentando correlação simples de $r = 0,19^*$ e correlação parcial de $r \text{ parc} = 0,20^*$, o que indica uma associação positiva e consistente. Esses resultados sugerem que o efeito do ACR sobre a capacidade de expansão não é apenas direto, mas pode refletir um comportamento integrado da arquitetura radicular que favorece a adaptação sob estresse hídrico.

Na safra de 2021, as estimativas de correlações fenotípicas simples e parciais entre os caracteres avaliados revelaram padrões distintos nas duas condições hídricas estudadas.

Tabela 9. Estimativas de correlações fenotípicas simples e parciais entre as características auxiliares e o rendimento de grãos (GY) e capacidade de expansão (PE) nas condições WW e WS, na safra de 2021.

Caráter	WW				WS			
	GY		PE		GY		PE	
	r	r parc	r	r parc	r	r parc	r	r parc
Auxiliar								
PH	-0,14	-0,26**	0,05	-0,05	-0,18*	-0,16	0,0	0,11
EH	0,08	0,2*	0,03	-0,01	0,01	0,1	-0,11	-0,09
SD	0,05	0,12	0,21**	0,24**	0,23**	0,26**	-0,16*	-0,15
CC	-0,17*	-0,16	0,21**	0,2*	0,01	-0,13	-0,03	-0,07
FLA	0,14	0,16	0,04	0,1	0,23**	0,3**	-0,05	0,04
ANT	0,04	0,01	-0,14	0,12	-0,08	-0,12	-0,12	-0,15
Fv/Fm	-0,12	-0,01	0,02	-0,02	-0,1*	-0,21*	-0,06	-0,02
ABR	-0,07	-0,15	-0,02	0,01	0,15	0,27**	0,07	0,03
NBR	-0,05	-0,07	0,1	0,23**	0,06	0,06	-0,06	-0,03
ACR	0,08	0,09	-0,13	-0,18*	-0,1	-0,23**	0,1	0,07
NCR	-0,05	-0,13	-0,2*	-0,35**	-0,04	-0,01	-0,04	-0,04
EL	0,23**	-0,01	0,07	0,09	0,34**	0,29**	0,07	-0,03
ED	0,38**	0,14	-0,06	-0,05	0,28**	0,07	0,16*	0,01
GNR	0,31**	0,22**	0,01	-0,12	0,27**	0,1	0,12	-0,1
RNE	0,49**	0,25**	0,0	-0,11	0,39**	0,23**	0,15	-0,08
GNE	0,47**	-0,2*	0,01	0,11	0,34**	-0,12	0,17*	0,11
100GW	0,27**	0,1	-0,07	-0,08	-0,01	-0,01	0,39**	0,32**

EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da

coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Novamente constata-se que associações podem ser significativas em uma determinada condição não o são em outra condição. De maneira geral as correlações parciais não indicaram mudanças de magnitude e sinal das associações simples obtidas. Assim, para este conjunto de variáveis, o impacto de outras variáveis sobre a associação de interesse foi pequeno.

Em relação às características agrônômicas (EL, ED, NGR, GNR, RNE, NGE, GW) e características morfológicas e fisiológicas (ANT, CC, FLA, Fv/Fm, SD, EH, PH, ABR, NBR, ACR, NCR), o grupo de características agrônômicas, especialmente aquelas relacionadas à espiga e ao enchimento de grãos, mostrou-se mais importante para explicar o rendimento de grãos. Variáveis como o número de fileiras por espiga (RNE), o número de grãos por espiga (GNE), o número de grãos por fileira (GNR), o comprimento (EL) e o diâmetro de espiga (ED) apresentaram correlações simples e parciais significativas com GY em ambas as condições hídricas. Por exemplo, em WW, RNE apresentou $r = 0,49^{**}$ e GNE $r = 0,47^{**}$, enquanto em WS o destaque permaneceu com RNE ($r = 0,39^{**}$) e EL ($r = 0,34^{**}$). Esses resultados confirmam que espigas bem desenvolvidas e com boa conformação estão diretamente associadas ao maior rendimento.

Por outro lado, para a capacidade de expansão (PE), as correlações foram, em geral, mais modestas e menos consistentes entre os grupos, com associações positivas esporádicas em variáveis como teor de clorofilas (CC) em WW ($r = 0,21^{**}$) ou massa de 100 grãos (100GW) em WS ($r = 0,39^{**}$). Esse padrão reforça que a expansão está mais ligada a atributos internos do grão, como composição de amido e estrutura do endosperma, do que a variáveis morfológicas tradicionais.

Dentro do grupo morfofisiológico, o diâmetro de colmo (SD) se destacou de forma particular. Em WW, SD apresentou correlações positivas e significativas com a capacidade de expansão (PE), tanto na análise simples ($r = 0,21^{**}$) quanto na parcial ($r_{\text{parc}} = 0,24^{**}$), evidenciando que plantas com colmos mais espessos tendem a produzir grãos com melhor qualidade. Já em WS, o SD manteve correlação positiva e significativa com o rendimento de grãos (GY), tanto na análise simples ($r = 0,23^{**}$) quanto na parcial ($r_{\text{parc}} = 0,26^{**}$), indicando que, mesmo sob estresse hídrico, o vigor estrutural da planta favorece a produção. Para PE em WS, entretanto, a correlação do SD foi negativa e significativa ($r = -0,16^*$), sugerindo que, sob condições adversas, o aumento no diâmetro de colmo não necessariamente

se traduz em maior capacidade de expansão, o que pode refletir um redirecionamento do uso de recursos da planta.

Esses resultados evidenciam que, entre os grupos avaliados, as características agronômicas continuam sendo as mais relevantes para o aumento do rendimento de grãos, independentemente da condição hídrica. Além disso, o diâmetro de colmo (SD) se destaca como um caráter auxiliar de interesse por demonstrar relações significativas tanto com GY quanto, em determinados contextos, com PE, funcionando como um indicador estrutural que pode ser utilizado em programas de melhoramento visando maior estabilidade produtiva e melhor qualidade dos grãos.

Sob condição bem irrigada (WW), observou-se que o número de grãos por fileira (GNR) apresentou correlações positivas com o rendimento ($r = 0,31$ e $r_{\text{parc}} = 0,22^{**}$), reforçando a importância desse componente como contribuinte direto para a produtividade. Já o número de grãos por espiga (GNE) também apresentou correlação significativa na análise simples ($r = 0,47^*$), porém com correlação parcial negativa ($r_{\text{parc}} = -0,20$, a 5%), apontando para uma relação mais complexa, na qual efeitos indiretos de outras características podem atenuar o impacto positivo inicialmente observado.

Em condição de estresse hídrico (WS), o rendimento de grãos foi positivamente associado ao diâmetro de colmo (SD), que apresentou $r = 0,23^{**}$ e $r_{\text{parc}} = 0,26^{**}$, evidenciando que colmos mais espessos favorecem a produtividade em ambientes com escassez. De forma semelhante, o comprimento de espigas (EL) mostrou correlações significativas ($r = 0,34^{**}$ e $r_{\text{parc}} = 0,29^{**}$), reforçando que espigas mais longas estão associadas a maior rendimento. O número de fileiras por espiga (RNE) também se destacou, com correlações positivas ($r = 0,39^{**}$ e $r_{\text{parc}} = 0,23^{**}$), caracterizando-se como um componente fundamental para a produção sob déficit hídrico. Além desses caracteres estruturais, observou-se que a fluorescência de clorofilas (F_v/F_m) apresentou correlação negativa significativa ($r = -0,10^*$ e $r_{\text{parc}} = -0,21^*$), possivelmente refletindo ajustes fisiológicos que, embora melhorem a eficiência fotoquímica, não se traduzem em maior rendimento. O teor relativo de flavonoides (FLA), por sua vez, mostrou correlação positiva com o rendimento ($r = 0,23^{**}$ e $r_{\text{parc}} = 0,30^{**}$), indicando que a maior produção de compostos antioxidantes pode ter favorecido a manutenção da produtividade em condições adversas. Por fim, o ângulo de raízes da coroa (ACR) apresentou correlação negativa significativa ($r = -0,10^*$ e $r_{\text{parc}} = -0,23^*$), sugerindo que arquiteturas radiculares menos favoráveis (ângulos mais abertos) podem ter limitado o aproveitamento de água e nutrientes, reduzindo a produção.

Quanto à capacidade de expansão dos grãos (PE), na condição WW, verificaram-se correlações positivas significativas com o diâmetro de colmo (SD) ($r = 0,21^{**}$ e $r_{\text{parc}} = 0,24^{**}$), indicando que plantas com colmos mais resistentes tendem a produzir grãos com melhor qualidade de expansão. Também houve correlações positivas entre PE e o teor relativo de clorofilas (CC) ($r = 0,21^*$ e $r_{\text{parc}} = 0,20^*$), sugerindo que plantas com maior capacidade fotossintética favorecem a qualidade do grão. Por outro lado, o número de raízes da coroa (NCR) apresentou correlações negativas com PE ($r = -0,20^{**}$ e $r_{\text{parc}} = -0,35^{**}$), indicando que arquiteturas radiculares podem comprometer a capacidade de expansão dos grãos.

Em condição WS, a característica que apresentou maior correlação significativa com a capacidade de expansão foi a massa de 100 grãos (100GW), com valores positivos ($r = 0,39^{**}$ e $r_{\text{parc}} = 0,32^{**}$), indicando que, sob déficit hídrico, grãos mais pesados tendem a apresentar melhor expansão, independentemente das demais características morfológicas ou fisiológicas.

4.3 Estudo dos Efeitos Totais, Diretos e Indiretos de Caracteres Auxiliares Sobre a Capacidade de Expansão e Rendimento de Grãos: Análise de trilha

Segundo Cruz et al. (2012), no contexto do melhoramento genético, é fundamental distinguir, entre as variáveis que apresentam correlação com o caráter de interesse, aquelas que exercem efeito direto positivo. Isso permite identificar características verdadeiramente associadas causalmente à variável principal, garantindo ganhos efetivos por meio da seleção. Com base nessa ideia, a análise de trilha foi utilizada para decompor as correlações das características morfológicas e agronômicas secundárias com o rendimento de grãos (GY) e com a capacidade de expansão (PE), utilizando dados da safra de 2020, sob condição hídrica bem irrigada (WW).

Na Tabela 10 são apresentadas as estimativas de efeitos diretos e indiretos de características morfológicas sobre o rendimento de grão no ensaio WW em 2020. Uma primeira interpretação diz respeito ao baixo coeficiente de determinação encontrado sendo indicativo de que o conjunto de características envolvidas no estudo não foi suficiente para descrever as variações no rendimento de grãos, nesta condição. Assim, infere-se sobre a necessidade de se ter outras informações mais relevantes ou, como será feito neste trabalho, apresentar estratégia de análise que permita avaliar, com o mesmo conjunto de dados, os efeitos de cada variável auxiliar usando abordagem que leva em consideração associações além das lineares utilizadas na presente análise.

Apesar do baixo coeficiente de determinação pode-se destacar, entre os caracteres morfológicos, o diâmetro do colmo (SD) que apresentou o maior efeito direto positivo (0,3607), sugerindo que plantas com colmos mais espessos tendem a ter maior suporte estrutural e potencial produtivo. Como este efeito é praticamente o mesmo do efeito total (correlação de Pearson, igual a 0,36) trata-se de uma relação de causa e efeito, não impactada por outra variável auxiliar mensurada.

O teor de flavonoides (FLAV) também teve contribuição direta relevante (0,2405) que também foi reflexo da correlação total igual a 0,21, possivelmente por refletir um metabolismo fisiológico eficiente. A altura da espiga (EH), com efeito direto de 0,2290, reforça a influência da arquitetura da planta na produtividade. O ângulo de raízes da coroa (ACR) apresentou um menor efeito direto (0,0608), mas indicou contribuição complementar ao desempenho produtivo.

Tabela 10. Efeitos diretos e indiretos de característica auxiliares sobre o rendimento de grãos (GY) sob condição hídrica bem irrigada (WW) em 2020.

Efeito de:	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	0,36	0,06	-0,01	-0,01	0,01	-0,03	-0,00	-0,00	0,01	0,00	0,00	0,38 **
EH	0,09	0,23	-0,12	-0,00	0,00	-0,01	0,00	0,00	-0,01	0,01	-0,00	0,20 *
PH	0,03	0,17	-0,16	-0,00	0,00	-0,01	-0,00	-0,00	-0,02	0,01	-0,00	0,02 ns
CC	0,13	0,03	-0,03	-0,03	0,00	0,02	-0,00	0,00	0,02	-0,00	0,00	0,13 ns
ANT	-0,09	-0,04	0,03	0,00	-0,03	-0,03	-0,00	-0,00	-0,04	-0,00	0,00	-0,20 *
FLAV	-0,04	-0,01	0,00	-0,00	0,00	0,24	-0,00	0,00	0,00	0,01	0,00	0,21 *
Fv/Fm	-0,04	0,00	0,00	0,00	0,00	-0,02	0,03	0,00	0,01	0,01	-0,00	0,01 ns
ABR	0,01	-0,00	-0,02	0,00	-0,00	-0,02	-0,00	-0,03	0,00	0,03	0,00	-0,04 ns
NBR	-0,02	0,03	-0,03	0,00	-0,01	-0,00	-0,00	0,00	-0,12	-0,01	0,00	-0,16 *
ACR	0,01	0,05	-0,01	0,00	0,00	0,03	0,00	-0,01	0,02	0,06	-0,00	0,13 ns
NCR	0,02	-0,02	0,00	-0,00	-0,01	0,06	-0,00	-0,00	-0,04	-0,01	0,01	-0,01 ns

$R^2 = 25,8\%$. $p_E = 0,86$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t

Na Tabela 11 foi considerado um outro conjunto de caracteres auxiliares, representado pelas características agronômicas, para explicar as variações do rendimento no

experimento WW, ano 2020. Novamente constata-se coeficiente de determinação relativamente baixo, igual a 27,07%.

Nas características agrônômicas, o número de grãos por fileira (NGR) destacou-se com efeito direto elevado (0,6971), seguido pelo diâmetro da espiga (ED) (0,1988) e pelo peso de 100 grãos (100GW) (0,1355). O comprimento da espiga (EL), mesmo com menor efeito direto (0,1217), demonstrou importância via efeitos indiretos, especialmente por GNR.

Tabela 11. Efeitos diretos e indiretos de características auxiliares sobre o rendimento de grãos (GY) sob condição hídrica bem irrigada (WW) em 2020

Efeito de	EL	ED	GNR	RNE	GNE	GW	Total
EL	0,12	-0,01	0,38	-0,07	0,12	0,039	0,34**
ED	-0,01	0,20	-0,06	0,11	-0,07	0,07	0,23**
GNR	0,07	-0,02	0,69	-0,03	-0,31	-0,00	0,41**
RNE	-0,04	0,11	-0,09	0,21	-0,16	-0,01	0,02
GNE	0,40	0,04	0,58	0,09	-0,37	-0,01	0,38**
GW	0,04	0,09	-0,02	-0,01	0,02	0,14	0,25**

$R^2 = 27,08\%$. $p_E = 0,85$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Os efeitos diretos e indiretos de dois conjuntos de caracteres auxiliares sobre a capacidade de expansão, nas condições bem irrigada (ww) em 2020, são apresentados nas Tabelas 12 e 13. De início destaca-se o fato do baixo coeficiente de determinação genotípica indicando que a abordagem deve ser vista com cautela por não levar em consideração outras variáveis de maior relevância, possivelmente incluída no grupo de variáveis residuais, ou de uma abordagem que contemple outros tipos de ações, além das lineares, entre as variáveis explicativas.

Apesar das limitações do baixo coeficiente de determinação, para a capacidade de expansão (PE), a característica morfológica com maior efeito direto na condição de bem irrigado foi o teor de antocianinas (ANT) (0,3882), possivelmente associado à estrutura dos grãos. Os caracteres auxiliares EH e NBR também se destacaram por apresentar correlações negativas e significativas associadas a efeitos diretos cujas magnitudes eram próximas aos efeitos totais.

Tabela 12. Efeitos diretos e indiretos de característica auxiliares sobre a capacidade de expansão (PE) sob condição de estresse hídrico (WW) em 2020

Efeito de	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	0.11	-0.06	0.01	-0.03	-0.10	-0.02	-0.00	-0.00	0.02	-0.00	-0.00	-0.08
EH	0.03	-0.24	0.12	-0.01	-0.07	-0.00	0.00	0.00	-0.04	-0.01	0.01	-0.21**
PH	0.01	-0.18	0.16	-0.02	-0.07	-0.00	-0.00	-0.01	-0.07	-0.00	0.00	-0.16*
CC	0.04	-0.03	0.03	-0.09	-0.06	0.01	-0.00	0.00	0.05	0.00	-0.00	-0.06
ANT	-0.03	0.04	-0.03	0.01	0.39	-0.01	-0.00	-0.00	-0.13	0.00	-0.01	0.23**
FLAV	-0.01	0.01	-0.00	-0.01	-0.04	0.13	-0.00	0.00	0.01	-0.00	-0.01	0.06
Fv/Fm	-0.01	-0.00	-0.00	0.02	-0.05	-0.01	0.02	0.01	0.04	-0.01	0.01	0.00
ABR	0.00	0.00	0.02	0.00	0.04	-0.01	-0.00	-0.04	0.00	-0.02	-0.00	-0.01
NBR	-0.01	-0.03	0.03	0.01	0.12	-0.00	-0.00	0.00	-0.40	0.00	-0.02	-0.28**
ACR	0.00	-0.05	0.01	0.00	-0.02	0.01	0.00	-0.02	0.06	-0.03	0.01	-0.01
NCR	0.00	0.03	-0.00	-0.01	0.10	0.03	-0.00	-0.00	-0.13	0.01	-0.05	-0.03

$R^2 = 23,13\%$. $p_E = 0,87$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre PE. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Entre os caracteres agronômicos, nas condições bem irrigada (WW) em 2020 (Tabela 13), apenas o número de grãos por espiga (GNE) mostrou efeito direto positivo (0,2693), apesar da correlação total negativa, devido a efeitos indiretos desfavoráveis. A característica ED foi a que apresentou melhor relação de causa e efeito com efeitos total e direto negativo, o que indica sua potencialidade para fins de seleção voltados à PE.

Tabela 13. Efeitos diretos e indiretos de característica auxiliares sobre a capacidade de expansão (PE) sob condição de estresse hídrico (WS) em 2020

Efeito de	EL	ED	GNR	RNE	GNE	GW	Total
EL	-0.03	0.03	-0.12	0.03	0.09	-0.00	-0.00
ED	0.00	-0.40	0.02	-0.04	0.05	-0.00	-0.37**
GNR	-0.02	0.04	-0.21	0.01	0.23	0.00	0.05
RNE	0.01	-0.22	0.03	-0.08	0.11	0.00	-0.15
GNE	-0.01	-0.08	-0.18	-0.03	0.27	0.00	-0.03
GW	-0.01	-0.20	0.01	0.00	-0.01	-0.01	-0.21**

$R^2 = 27\%$. $p_E = 0,83$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre PE. EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste de Pearson.

Em condição de estresse hídrico (WS) na safra de 2020, identificamos características com efeitos diretos positivos sobre GY e PE.

Nas Tabelas 14 e 15 são mostrados os efeitos diretos e indiretos de característica auxiliares sobre o rendimento de grãos (GY) sob condição de déficit hídrico (WS) em 2020. Em um primeiro momento, chama atenção o baixo coeficiente de determinação, o que sugere que os resultados devem ser interpretados com cautela. Isso porque a análise pode não ter considerado variáveis de maior relevância ou ainda pode demandar uma abordagem que inclua outros tipos de relações, para além das puramente lineares, entre as variáveis explicativas.

Mesmo considerando as limitações associadas ao baixo coeficiente de determinação, entre os caracteres morfológicos (Tabela 14), o teor relativo a clorofila (CC) destacou-se, sob estresse, com elevado efeito direto 0,3356, seguido pelo número de raízes de suporte (NBR) (0,2627), o ângulo de raízes da coroa (ACR) (0,1819) e o diâmetro do colmo (SD) (0,1643) também apresentaram efeitos diretos relevantes.

Tabela 14. Efeitos diretos e indiretos de característica auxiliares sobre o rendimento de grãos (GY) sob condição hídrica bem irrigada (WS) em 2020.

	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	0.16	-0.03	-0.01	0.08	0.0	-0.01	0.0	-0.0	0.05	0.02	-0.0	0.27**
EH	0.04	-0.12	0.05	0.05	-0.03	0.01	-0.0	-0.0	0.04	0.03	-0.01	0.07
PH	-0.02	-0.09	0.08	0.0	-0.03	0.01	-0.0	-0.0	0.01	0.02	-0.0	-0.02
CC	0.04	-0.02	0.0	0.34	0.03	0.0	-0.0	-0.0	-0.06	0.02	-0.01	0.33**
ANT	-0.01	-0.05	0.03	-0.12	-0.09	0.01	0.0	0.0	0.07	0.01	-0.0	-0.15
FLAV	-0.01	-0.02	0.01	0.01	-0.0	0.09	-0.01	-0.0	-0.02	0.03	-0.01	0.07
Fv/Fm	0.04	0.01	-0.0	-0.06	-0.01	-0.02	0.02	0.01	0.04	0.01	-0.0	0.03
ABR	0.02	-0.01	0.0	0.04	0.0	0.01	-0.0	-0.04	-0.04	0.04	-0.0	0.01
NBR	0.03	-0.02	0.0	-0.08	-0.02	-0.01	0.0	0.01	0.26	-0.03	-0.01	0.14
ACR	0.02	-0.02	0.01	0.03	-0.01	0.01	0.0	-0.01	-0.05	0.18	0.01	0.18*
NCR	0.01	-0.02	0.0	0.07	-0.01	0.02	0.0	-0.0	0.05	-0.03	-0.04	0.05

$R^2 = 23,18\%$. $p_E = 0,87$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Para os caracteres agronômicos em GY, sob estresse (tabela 15), o número de grãos por espiga (GNE) teve o maior efeito direto (1,7583), seguido por EL (0,2030), ED(0,0943). GNR e RNE apresentaram efeitos diretos negativos, apesar das correlações totais positivas.

Tabela 15. Efeitos diretos e indiretos de características auxiliares sobre o rendimento de grãos (GY) sob condição de estresse hídrico (WS) em 2020

	EL	ED	GNR	RNE	GNE	GW	Total
EL	0.20	0.01	-0.73	0.07	0.70	0.01	0.26**
ED	0.02	0.09	-0.11	-0.44	0.65	0.04	0.26**
GNR	0.13	0.01	-1.14	-0.08	1.37	-0.02	0.27**
RNE	-0.02	0.05	-0.10	-0.91	1.20	-0.00	0.22**
GNE	0.08	0.03	-0.89	-0.62	1.76	-0.01	0.35**
GW	0.02	0.04	0.17	0.04	-0.24	0.11	0.14

$R^2 = 19,53\%$. $p_E = 0,89$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Para a capacidade de expansão, ainda na safra de 2020, na condição WS (tabela 16), observa-se que o coeficiente de determinação genotípica obtido foi relativamente baixo ($R^2 = 27,65\%$). Esse resultado indica que o conjunto de características avaliadas explica apenas uma fração da variação em PE. Contudo, seguimos com cautela a análise. A variável Fv/Fm destacou-se por apresentar um efeito total positivo (0,29**), tornando-se a variável de maior correlação total positiva sobre a capacidade de expansão, com efeito direto de 0,2153. Esse resultado sugere que genótipos com maior eficiência fotoquímica tendem a apresentar melhor capacidade de expansão mesmo sob estresse hídrico. Em seguida temos ACR que apresentou efeito direto positivo (0,20) e efeito total igualmente positivo e muito próximo (0,19*), evidenciando que sua influência sobre a capacidade de expansão é predominantemente direta, com efeitos indiretos praticamente nulos ou compensados. Isso indica que a seleção baseada em ACR tende a resultar em ganhos consistentes para PE, tornando-se uma característica auxiliar de interesse em ambientes sob estresse hídrico. As demais variáveis morfológicas apresentaram efeitos diretos negativos.

Tabela 16. Efeitos diretos e indiretos de característica auxiliares sobre a capacidade de expansão (PE) sob condição de estresse hídrico (WS) em 2020

	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	0,06	-0,02	0,01	-0,04	0,01	0,02	0,05	0,00	-0,02	0,02	-0,00	0,09
EH	0,01	-0,06	-0,05	-0,02	-0,04	-0,03	-0,02	0,00	-0,02	0,03	-0,02	-0,22**
PH	-0,01	-0,04	-0,08	-0,00	-0,04	-0,03	-0,00	0,00	-0,01	0,02	-0,01	-0,20*
CC	0,01	-0,01	-0,00	-0,16	0,04	-0,01	-0,04	0,00	0,02	0,02	-0,02	-0,14
ANT	-0,00	-0,02	-0,03	0,06	-0,11	-0,01	0,02	0,00	-0,02	0,01	-0,01	-0,12
FLAV	-0,00	-0,01	-0,01	-0,00	-0,01	-0,22	-0,06	0,00	0,01	0,03	-0,03	-0,30*
Fv/Fm	0,01	0,01	0,00	0,03	-0,01	0,06	0,21	-0,00	-0,01	0,01	-0,01	0,29**
ABR	0,01	-0,00	-0,00	-0,02	0,00	-0,01	-0,04	0,01	0,01	0,04	-0,00	-0,00
NBR	0,01	-0,01	-0,00	0,04	-0,03	0,02	0,03	-0,00	-0,10	-0,03	-0,02	-0,10
ACR	0,01	-0,01	-0,01	-0,02	-0,01	-0,03	0,01	0,00	0,02	0,20	0,02	0,19*
NCR	0,00	-0,01	-0,00	-0,03	-0,01	-0,06	0,01	0,00	-0,02	-0,03	-0,12	-0,27**

$R^2 = 27,65\%$. $p_E = 0,85$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre PE. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Nas Tabelas 17 a 22 são apresentados os resultados dos efeitos diretos e indiretos sobre GY e PE, em diferentes condições de fornecimento de água, no experimento realizado em 2021. De maneira geral, os resultados demonstram que diferentes conjuntos de características exercem influências distintas sobre GY e PE.

É importante notarmos que o coeficiente de determinação foi baixo ($R^2 = 0,27$) o que indica que as variáveis analisadas explicam apenas uma parte limitada do rendimento de grãos sob estresse hídrico. Isso sugere cautela na interpretação, pois outras características relevantes podem não ter sido incluídas no modelo, além de o método considerar apenas relações lineares.

Contudo, mesmo com tais limitações e um baixo coeficiente, para a safra de 2021 sob condição bem irrigada (WW), nenhuma característica morfofisiológica apresentou efeito total positivo e significativo sobre o rendimento de grãos (GY). O único efeito total significativo observado foi negativo, associado ao teor relativo de clorofila de -0,17.

Para as variáveis agrônômicas (Tabela 17), da mesma safra, o número de fileiras por espiga (RNE) teve destaque com um efeito total de 0,49, com efeito direto de 1,65. Outro destaque vem de GNE, mesmo com efeito direto negativo (-2,53), demonstrou importância via efeitos indiretos, principalmente por GNR e RNE. Ainda que com efeitos diretos pouco

expressivo seguimos por ED (0,38**) - baixo efeito direto, mas relevante via efeitos indiretos, GNR (0,31**), GW(0,27**) e EL (0,23**).

Tabela 17. Efeitos diretos e indiretos de característica auxiliares sobre o rendimento de grão sob condição bem irrigada (WW) em 2021

	EL	ED	GNR	RNE	GNE	GW	Total
EL	-0.01	0.03	1.22	0.37	-1.40	0.03	0.23**
ED	-0.00	0.15	0.18	0.81	-0.81	0.04	0.38**
GNR	-0.01	0.01	2.05	0.42	-2.19	0.02	0.31**
RNE	-0.00	0.07	0.53	1.65	-1.78	0.03	0.49**
GNE	-0.01	0.05	1.77	1.16	-2.53	0.03	0.47**
GW	-0.00	0.07	0.38	0.45	-0.72	0.10	0.27**

$R^2 = 33,56\%$. $p_E = 0,81$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Na análise para PE em 2021, sob WW (tabela 18), O coeficiente de determinação foi baixo ($R^2 = 0,22$) indicando que as variáveis avaliadas explicam apenas parte da variação na capacidade de expansão (PE), o que exige cautela na interpretação dos resultados. Mesmo assim, os efeitos totais mostram associações relevantes: a altura de plantas (PH) apresentou efeito total de 0,21, o teor relativo de clorofilas (CC) também teve efeito total de 0,21 .

Ao analisar os efeitos diretos, a PH destacou-se com 0,28, sendo a maior contribuição direta positiva para PE, seguida de CC, com efeito direto de 0,20, ambos indicando características favoráveis à seleção indireta. Em contrapartida, o NCR apresentou efeito direto negativo de -0,42, reforçando sua associação desfavorável à capacidade de expansão.

Tabela 18. Efeitos diretos e indiretos de características auxiliares sobre a capacidade de expansão (PE) sob condição bem irrigada (WW) em 2021

	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	-0,06	-0,01	0,06	0,00	-0,00	-0,01	-0,00	0,00	-0,04	0,01	0,11	0,05
EH	-0,05	-0,02	0,08	-0,00	0,01	-0,01	-0,00	0,00	0,00	-0,01	0,02	0,03
PH	-0,01	-0,01	0,28	0,02	-0,00	-0,06	0,00	0,00	0,04	0,01	-0,07	0,21**
CC	-0,00	0,00	0,03	0,20	0,00	-0,04	-0,01	0,00	-0,02	0,00	0,04	0,21**
ANT	0,00	-0,00	-0,01	0,00	0,09	0,00	-0,00	0,00	-0,01	0,01	-0,06	0,04
FLAV	0,00	0,00	-0,13	-0,06	0,00	0,14	-0,00	0,00	-0,02	-0,08	0,01	-0,14
Fv/Fm	-0,02	-0,00	-0,04	0,07	0,00	0,01	-0,02	0,00	-0,03	-0,00	0,05	0,02
ABR	-0,00	-0,00	0,05	0,02	0,02	0,01	-0,00	0,01	-0,02	-0,11	0,01	-0,02
NBR	0,01	-0,00	0,05	-0,01	-0,00	-0,01	0,00	-0,00	0,24	0,02	-0,20	0,10
ACR	0,00	-0,00	-0,02	-0,00	-0,01	0,05	-0,00	0,01	-0,02	-0,22	0,09	-0,13
NCR	0,02	0,00	0,04	-0,02	0,01	-0,00	0,00	-0,00	0,12	0,05	-0,42	-0,20*

$R^2 = 21,92\%$. $p_E = 0,88$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre PE. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Para as variáveis agrônômicas avaliadas (safra 2021, condição WW), o coeficiente de determinação foi extremamente baixo ($R^2 = 0,03$), com efeito residual elevado (98,46%). Esse resultado indica que o conjunto de características analisadas praticamente não explica a variação observada em PE, o que inviabiliza interpretações consistentes.

Além disso, nenhum efeito total ou direto foi estatisticamente significativo, reforçando que, mesmo com cautela, não há elementos suficientes para apontar relações úteis ou tendências seguras. Diante disso, opta-se por não realizar uma análise detalhada para essas variáveis, já que os dados disponíveis não permitem extrair conclusões confiáveis.

Sob WS em 2021(tabela 19), entendendo a limitação por conta de um baixo coeficiente de determinação (0,28), em relação a características morfológicas, temos em destaque o teor de antocianina (ANT) (0,32), seguida do PH (0,27) que apresentaram os maiores efeitos diretos sobre GY.

Tabela 19. Efeitos diretos e indiretos de característica auxiliares sobre o rendimento de grão em condição de estresse hídrico (WS) em 2021

	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	-0,25	0,11	0,05	-0,00	-0,00	0,01	-0,04	-0,09	0,01	0,03	-0,00	-0,18*
EH	-0,18	0,15	0,07	0,01	0,07	-0,01	-0,07	-0,05	0,01	0,02	-0,00	0,01
PH	-0,05	0,04	0,27	-0,03	0,01	0,04	-0,03	-0,03	0,00	0,02	-0,00	0,23**
CC	-0,01	-0,01	0,07	-0,13	0,06	0,02	0,02	0,05	-0,01	-0,04	-0,00	0,01
ANT	0,00	0,03	0,00	-0,02	0,32	-0,02	-0,04	-0,01	0,00	-0,02	-0,00	0,23**
FLAV	0,02	0,02	-0,09	0,03	0,06	-0,11	-0,00	0,02	0,00	-0,01	-0,00	-0,08
Fv/Fm	-0,05	0,05	0,04	0,01	0,06	-0,00	-0,21	-0,02	-0,01	0,02	0,00	-0,10
ABR	0,07	-0,03	-0,03	-0,02	-0,02	-0,01	0,01	0,29	-0,02	-0,11	0,00	0,15
NBR	-0,05	0,02	0,01	0,03	0,01	-0,00	0,03	-0,11	0,06	0,07	-0,00	0,06
ACR	0,03	-0,01	-0,02	-0,02	0,03	-0,01	0,02	0,13	-0,02	-0,24	-0,00	-0,10
NCR	-0,10	0,05	0,01	-0,01	0,07	-0,01	0,00	-0,05	0,01	-0,01	-0,01	-0,04

$R^2 = 27,95$. $p_E = 0,84$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

No grupo agrônomico (Tabela 20), mesmo com um R^2 relativamente baixo, pode-se notar que variáveis ligadas ao número de grãos e ao tamanho das espigas exercem influência direta e positiva sobre o rendimento de grãos em condição de estresse hídrico. O número de fileiras por espiga (RNE) apresentou o maior efeito direto (0,76), destacando-se como a característica de maior potencial para promover ganhos em GY. Em seguida, aparecem o número de grãos por fileira (GNR) com efeito direto de 0,60 e o comprimento de espigas (EL) com 0,37, todas com contribuições diretas relevantes e significativas, configurando-se como variáveis auxiliares recomendadas para a seleção indireta visando maior rendimento. Observa-se também que GNE, mais uma vez, com efeito direto negativo (-1,06), demonstrou importância via efeitos indiretos, principalmente por GNR e RNE.

Tabela 20. Efeitos diretos e indiretos de característica auxiliares sobre o rendimento de grão sob condição em déficit hídrico (WS) em 2021

	EL	ED	GNR	RNE	GNE	GW	Total
EL	0,37	0,03	0,37	0,18	-0,61	0,00	0,34**
ED	0,16	0,08	0,25	0,31	-0,52	-0,00	0,28**
GNR	0,23	0,03	0,60	0,41	-1,00	-0,00	0,27**
RNE	0,09	0,03	0,32	0,76	-0,82	-0,00	0,39**
GNE	0,21	0,04	0,56	0,59	-1,06	-0,00	0,34**
GW	0,00	0,02	0,01	0,07	-0,10	-0,01	-0,01

$R^2 = 24,54\%$. $p_E = 0,87$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre GY. EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Para PE sob WS em 2021 (Tabela 21), os efeitos foram menores, evidenciado por coeficiente de determinação muito baixo ($R^2 = 0,08$), o que significa que as variáveis avaliadas explicam apenas uma pequena fração da variação observada na capacidade de expansão. Isso implica que a maior parte da variação permanece nos resíduos, já que outras características importantes para explicar a PE não foram abordadas ou as relações podem não ser apenas lineares.

Ao observar os efeitos totais percebe-se que a grande maioria dos valores está próxima de zero e não é significativa. Isso indica baixa associação linear das variáveis estudadas com a capacidade de expansão. O único efeito total que apresentou significância estatística foi o da altura de plantas (PH), com um efeito total de $-0,16$, indicando uma associação negativa entre plantas mais altas e a capacidade de expansão, mesmo assim com efeito de magnitude reduzida.

Tabela 21. Efeitos diretos e indiretos de característica auxiliares sobre a capacidade de expansão em condição de estresse hídrico (WS) em 2021

	SD	EH	PH	CC	ANT	FLA	Fv/Fm	ABR	NBR	ACR	NCR	Total
SD	0,20	-0,12	-0,03	-0,00	-0,00	0,01	-0,01	-0,01	-0,01	-0,01	-0,02	0,00
EH	0,15	-0,16	-0,05	0,01	0,01	-0,02	-0,01	-0,01	-0,00	-0,01	-0,01	-0,11
PH	0,04	-0,04	-0,17	-0,02	0,00	0,06	-0,00	-0,00	-0,00	-0,01	-0,00	-0,16*
CC	0,00	0,01	-0,04	-0,08	0,01	0,04	0,00	0,01	0,01	0,01	-0,00	-0,03
ANT	-0,00	-0,03	-0,00	-0,01	0,04	-0,03	-0,00	-0,00	-0,00	0,01	-0,01	-0,05
FLAV	-0,01	-0,02	0,06	0,02	0,01	-0,17	-0,00	0,00	-0,00	0,00	-0,00	-0,12
Fv/Fm	0,04	-0,05	-0,03	0,01	0,01	-0,00	-0,02	-0,00	0,00	-0,01	0,00	-0,06
ABR	-0,06	0,03	0,02	-0,01	-0,00	-0,01	0,00	0,04	0,01	0,04	0,01	0,07
NBR	0,04	-0,02	-0,00	0,02	0,00	-0,00	0,00	-0,01	-0,04	-0,03	-0,01	-0,06
ACR	-0,02	0,02	0,01	-0,01	0,00	-0,01	0,00	0,02	0,01	0,08	-0,00	0,10
NCR	0,08	-0,05	-0,00	-0,01	0,01	-0,01	0,00	-0,01	-0,01	0,00	-0,05	-0,04

$R^2 = 8,36\%$. $p_E = 0,96$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre PE. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; ; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

A análise agronômica sobre PE (Tabela 22), apresentou um coeficiente de determinação baixo ($R^2 = 0,18$), indicando que as variáveis analisadas explicam apenas uma pequena parcela da variação na capacidade de expansão (PE) sob condição de déficit hídrico em 2021. Ainda assim, alguns resultados merecem destaque.

Entre as variáveis avaliadas, a massa de 100 grãos (GW) se sobressaiu, apresentando o maior efeito direto positivo (0,33) e o maior efeito total (0,39). Isso evidencia que genótipos com maior massa de grãos tendem a expressar maior capacidade de expansão mesmo em condição de estresse hídrico, configurando-se como a principal característica auxiliar para seleção indireta.

Além disso, o diâmetro de espigas (ED) apresentou efeito total positivo de 0,16, no entanto, seu efeito direto foi baixo (ED = 0,01), indicando que a influência favorável dessas variáveis ocorre principalmente por efeitos indiretos via outras características. O número de grãos por espiga (GNE) obteve efeito total de 0,17 e um alto efeito direto, o que indica a importância dessa variável para a seleção.

Tabela 22. Efeitos diretos e indiretos de característica auxiliares sobre a capacidade de expansão sob condição em déficit hídrico (WS) em 2021

	EL	ED	GNR	RNE	GNE	GW	Total
EL	-0,04	0,01	-0,39	-0,07	0,56	0,00	0,07
ED	-0,02	0,01	-0,26	-0,12	0,48	0,06	0,16*
GNR	-0,02	0,01	-0,63	-0,16	0,92	0,01	0,12
RNE	-0,01	0,01	-0,34	-0,30	0,75	0,03	0,15
GNE	-0,02	0,01	-0,60	-0,23	0,98	0,03	0,17*
GW	-0,00	0,00	-0,01	-0,03	0,09	0,33	0,39**

$R^2 = 17,83\%$. $p_E = 0,91$

(*) Diagonal: representa o efeito direto da característica auxiliar. Na linha e fora da diagonal representa o efeito indireto da variável da linha, via a variável da coluna, sobre PE. EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

Ao analisar os coeficientes de determinação (R^2) obtidos nas diferentes condições hídricas e safras, é possível refletir com mais profundidade sobre os limites e as potencialidades da análise de trilha dentro do contexto do melhoramento genético do milho-pipoca.

Como apontado por Cruz et al. (2012), o coeficiente de determinação expressa o quanto da variação da característica principal — como rendimento de grãos ou capacidade de expansão — é explicada pelas variáveis consideradas no modelo. De forma simples, quanto

maior esse valor, mais confiável tende a ser a interpretação dos efeitos diretos e indiretos identificados. No entanto, quando o R^2 é baixo, isso pode indicar que variáveis importantes estão sendo deixadas de fora ou que as relações entre os dados são mais complexas do que o modelo consegue captar.

Nos resultados obtidos, observamos uma grande oscilação nesses valores. Em alguns cenários, como o rendimento de grãos sob condição bem irrigada em 2021, o modelo com variáveis agronômicas alcançou um R^2 de 0,34, o que significa que aproximadamente 34% da variação do rendimento foi explicada. É um valor razoável, que dá boas pistas para a seleção. No entanto, em outros casos, como a capacidade de expansão com irrigação adequada no mesmo ano, o coeficiente foi de apenas 0,03. Em outras palavras, só 3% da variação foi explicada, o que nos alerta sobre a limitação real da análise de trilha nesse tipo de situação.

Esses números sugerem algo importante: nem sempre a análise de trilha consegue captar a complexidade das relações envolvidas no desempenho das plantas — especialmente em ambientes contrastantes. Isso se deve, em grande parte, à suposição de que todas as relações entre as variáveis são lineares. No entanto, em sistemas biológicos, as interações entre características raramente seguem uma lógica tão simples. Como destacam Garcia-Campos et al. (2015), quando há muitas variáveis interagindo, como ocorre em condições de estresse, a linearidade deixa de representar bem a realidade, e o modelo pode acabar deixando escapar informações valiosas. Além disso, conforme lembra Klem (1995), os resultados da trilha podem não se manter consistentes em outros ambientes, o que exige cautela ao extrapolar conclusões.

Diante disso, é fundamental reconhecer que, embora continue sendo uma ferramenta útil e amplamente usada, a análise de trilha não deve ser empregada de forma isolada. Existem hoje metodologias mais flexíveis, como as árvores de decisão e algoritmos de aprendizado de máquina, que têm mostrado bom desempenho na análise de dados complexos. Mesmo que esses métodos não tenham sido originalmente desenvolvidos para esse tipo de análise, eles podem ser adaptados para ajudar a identificar padrões não lineares e interações difíceis de serem visualizadas em modelos tradicionais.

4.4 Estudo dos Efeitos Totais, Diretos e Indiretos de Caracteres Auxiliares Sobre a Capacidade de Expansão e Rendimento de Grãos: Árvore de Decisão

4.4.1. Eficiência de predição de características principais por meio de características auxiliares

Nesta seção, são apresentados os resultados obtidos por meio da aplicação de árvores de decisão nas duas safras avaliadas (2020 e 2021).

Na Tabela 23 é apresentada a comparação entre os valores de eficiência de predição obtidos a partir de modelos lineares (análise de trilha) e não lineares (árvore de decisão) em diferentes cenários e condições experimentais. De maneira geral, observa-se que, em todos os cenários avaliados (WW 2020, WS 2020, WW 2021 e WS 2021), as árvores de decisão apresentaram valores de eficiência de predição consistentemente superiores aos obtidos pela análise de trilha. Neste ponto, deve-se destacar que o presente trabalho não buscou modelos que maximizassem a capacidade preditiva das características auxiliares e, portanto, a eficácia dos refinamentos das árvores de decisão, tais como os procedimentos de bagging, random forest ou boosting, não foram investigados. A ênfase na abordagem fundamentada na árvore de decisão se dá em razão dela apresentar, como resultado principal, topologias cujo conteúdo e bifurcações permitem uma interpretação adequada de quantas e quais variáveis auxiliares poderiam ser utilizadas em trabalhos futuros de melhoramento. Assim, considerou-se que, para este fim que é a escolha de características auxiliares, a árvore de decisão é a abordagem que melhor contrastaria com a tradicional representada pela análise de trilha.

Comparação da eficiência de predição entre modelos lineares (análise de trilha) e não-lineares (árvore de decisão)

Cenário	WW 2020		WS 2020		WW2021		WS2021	
	Análise de Trilha	Árvore de Dec.	Análise de Trilha	Árvore de Dec.	Análise de Trilha	Árvore de Dec.	Análise de Trilha	Árvore de Dec.
1	0,25	0,60	0,23	0,40	0,17	0,49	0,28	0,55
2	0,23	0,69	0,28	0,60	0,22	0,49	0,08	0,53
3	0,27	0,51	0,19	0,53	0,34	0,61	0,25	0,62
4	0,15	0,47	0,03	0,39	0,03	0,44	0,18	0,51

1: Características Morfológicas e Rendimento de Grão

2: Características Morfológicas e Capacidade de Expansão

3: Características Agronômicas e Rendimento de Grão

4: Características Agronômicas e Capacidade de Expansão

Como destaque ressalta-se que no cenário, envolvendo características morfológicas e rendimento de grãos (cenário 1), a análise de trilha apresentou eficiências variando de 0,17 a 0,28, enquanto as árvores de decisão alcançaram valores entre 0,40 a 0,60, independentemente da condição hídrica e do ano de avaliação. De forma semelhante, no cenário características morfológicas e capacidade de expansão (cenário 2), os valores da análise de trilha oscilaram

entre 0,08 e 0,28, ao passo que as árvores de decisão variaram de 0,49 a 0,69. Essa tendência se manteve para os cenários envolvendo características agronômicas: no rendimento de grãos (cenário 3), a análise de trilha apresentou eficiências entre 0,19 e 0,34, enquanto as árvores de decisão atingiram 0,51 a 0,62. Já na capacidade de expansão (cenário 4), os valores foram de 0,03 a 0,18 para a análise de trilha, contra 0,39 a 0,51 para as árvores de decisão.

Os resultados obtidos deixam claro que as árvores de decisão têm uma habilidade maior para capturar relações complexas e não lineares entre as variáveis, o que se traduz em um poder preditivo superior. Um exemplo disso é o estudo de Chauhan et al. (2024), no qual as árvores de decisão alcançaram 83,5% de precisão na previsão da produtividade da safra, superando os 81,2% obtidos pela regressão linear.

Essa diferença acontece porque a análise de trilha, por ser um método que mede associação linear, parte do pressuposto de que as relações entre as variáveis são aditivas e independentes. Com isso, ela pode deixar de captar interações ou efeitos condicionais que, na realidade, influenciam de forma significativa o caráter principal em estudo. Já as árvores de decisão, como destacam Piyar e Baulkani (2024), são eficazes para lidar com relações e interações não lineares entre vários atributos, algo muito comum em dados agrícolas, que envolvem características interligadas do solo, do clima e da cultura.

Como salientado por Cruz et al. (2012), o R^2 expressa o quanto da variação de uma característica principal — como rendimento de grãos ou capacidade de expansão — é explicada pelas variáveis incluídas no modelo. Valores baixos podem indicar que variáveis importantes foram omitidas ou que as relações entre os dados são mais complexas do que o modelo linear é capaz de captar. Nesses contextos, as árvores de decisão apresentam-se como ferramentas superiores, pois conseguem lidar com relações não lineares e fornecer caminhos preditivos mais alinhados à complexidade biológica dos caracteres estudados.

4.4.2. Árvores de decisão

Em razão da grande quantidade de combinações entre ano, condição hídrica e variável-resposta (rendimento de grãos e capacidade de expansão), optou-se por destacar, no corpo do texto, uma safra e uma característica (morfológica ou agronômica) em ambas as condições hídricas. Os demais gráficos, igualmente relevantes, foram organizadas no APÊNDICE B e estão disponíveis para consulta.

A árvore de decisão ajustada (Figura 3) com base nos dados morfofisiológicos da safra de 2020, sob condição bem irrigada (WW), identificou cinco variáveis explicativas

como relevantes para o rendimento de grãos (GY): diâmetro do colmo (SD), ângulo de raízes da coroa (ACR), número de raízes de suporte (NBR), teor relativo de flavonoides (FLA) e teor relativo de clorofilas (CC). Dentre essas, a variável SD foi fixada na raiz da árvore, evidenciando-se como a mais importante, o que indica grande associação com RG. A topologia da árvore permitiu a identificação de combinações hierárquicas organizadas entre variáveis morfológicas e fisiológicas. Após a poda, a árvore foi reduzida a sete nós terminais, mantendo, ainda assim, desempenho estatístico satisfatório, com coeficiente de determinação (R^2) igual a 0,57, correlação entre valores observados e preditos de 0,75. Tais resultados reiteram o potencial das árvores de decisão como ferramenta complementar na identificação de características secundárias relevantes, contribuindo de forma estratégica para programas de melhoramento.

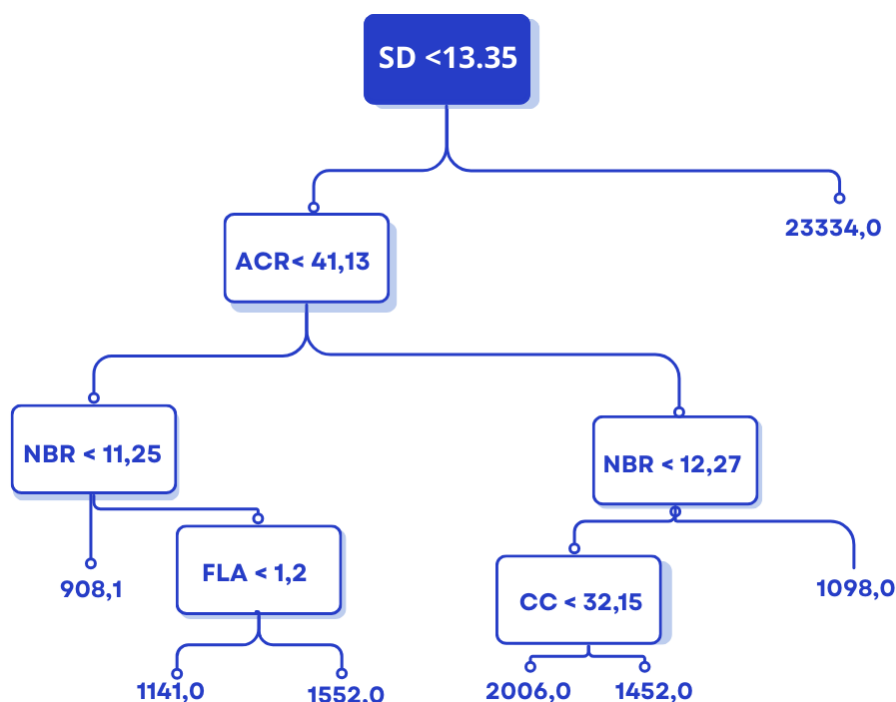


Figura 3 - Árvore de decisão para o rendimento de grãos (RG) em milho-pipoca, safra de 2020, sob condição bem irrigada (WW). SD – Diâmetro de colmo; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; FLA – Teor relativo de flavonoides; CC – Teor relativo de clorofilas. Os valores nos nós terminais representam as médias preditas de rendimento de grãos ($\text{kg} \cdot \text{ha}^{-1}$).

Por outro lado, a árvore de decisão gerada para os dados da mesma safra sob condição de estresse hídrico (WS) revelou uma estrutura distinta (Figura 4), o teor relativo de clorofilas (CC) foi fixado na raiz da árvore, indicando sua maior contribuição inicial na segmentação dos dados e, conseqüentemente, sua elevada associação com o desempenho produtivo. Na

sequência hierárquica, participaram da divisão o diâmetro do colmo (SD) — que apareceu em dois pontos distintos da árvore —, o número de raízes de suporte (NBR) e a variável fluorescência de clorofilas Fv/Fm.

Após a poda, a árvore apresentou nove nós terminais, com valores preditos de rendimento variando entre 223,29 e 699,61 kg/ha, refletindo a diversidade de respostas entre os genótipos avaliados. Ainda que não tenha apresentado um alto coeficiente de correlação ($R^2=0,40$), a topologia da árvore forneceu informações relevantes sobre a importância das características secundárias sob estresse.

A presença de variáveis fisiológicas em posições estratégicas da árvore de decisão sob condição WS reforça o papel adaptativo dessas características em ambientes adversos. O teor relativo de clorofilas (CC), por exemplo, foi fixado na raiz da árvore, destacando-se como o traço mais discriminativo entre os genótipos avaliados. Tal comportamento está em consonância com Kamphorst et al. (2021), que identificaram o CC como uma variável fortemente associada ao rendimento de grãos em milho-pipoca submetido ao déficit hídrico, recomendando seu uso como critério auxiliar na seleção indireta. Da mesma forma, a fluorescência de clorofilas, também presente na estrutura decisória da árvore, tem sido amplamente reconhecido como indicador de eficiência fotossintética e resposta ao estresse (OLIVEIRA et al., 2024). Além disso, o reaparecimento do diâmetro do colmo (SD) em diferentes níveis da árvore indica sua estabilidade como critério de seleção.

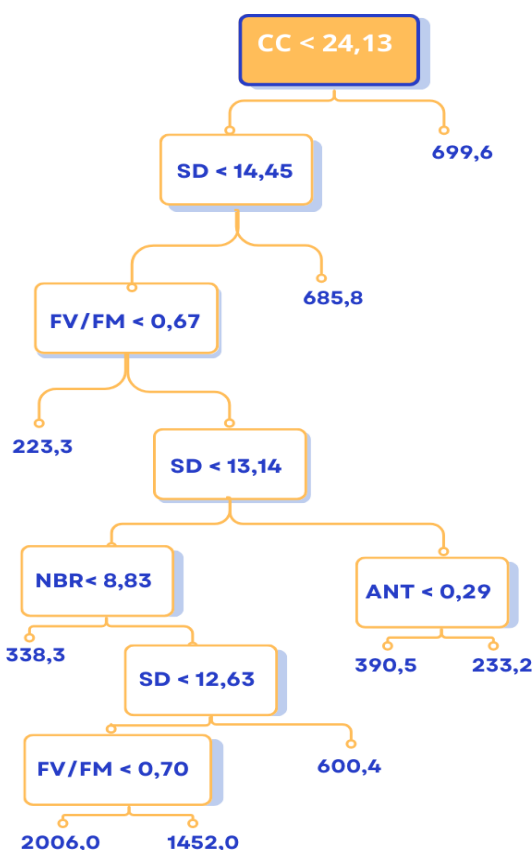


Figura 4 - Árvore de decisão para o rendimento de grãos (RG) em milho-pipoca, safra de 2020, sob condição hídrica WS. CC – Teor relativo de clorofilas; SD – Diâmetro de colmo; F_v/F_m – Fluorescência de clorofilas; NBR – Número de raízes de suporte; ANT – Teor relativo de antocianinas. Os valores nos nós terminais representam as médias previstas de rendimento de grãos (kg.ha⁻¹).

Para safra de 2021, a análise da árvore de decisão construída com base nas variáveis agronômicas, sob condição bem irrigada (WW), permitiu identificar padrões importantes na explicação da capacidade de expansão (Figura 5). Entre as variáveis consideradas, destacaram-se o comprimento e o diâmetro das espigas (EL e ED), o número de grãos por fileira (GNR), o número de fileiras por espiga (RNE), o total de grãos por espiga (GNE) e o peso de 100 grãos (100GW).

Pela estrutura da árvore fica evidente que o peso de 100 grãos (100GW) é o principal divisor inicial, indicando que genótipos com menor massa tendem a apresentar maior capacidade de expansão. Esse comportamento também foi observado por Song et al. (1991), que relataram uma associação entre menor massa e densidade dos grãos com maior capacidade de expansão volumétrica. Além disso, características como o comprimento das espigas (EL) e o número de fileiras (RNE) surgiram em divisões subsequentes, sugerindo que

sua interação com o 100GW contribui para formar grupos com diferentes níveis de desempenho quanto à PE.

Mesmo após o processo de poda, a árvore manteve uma estrutura clara e informativa, apresentando desempenho estatístico satisfatório ($R^2 = 0,44$; correlação = 0,66; RMSE = 3,28).

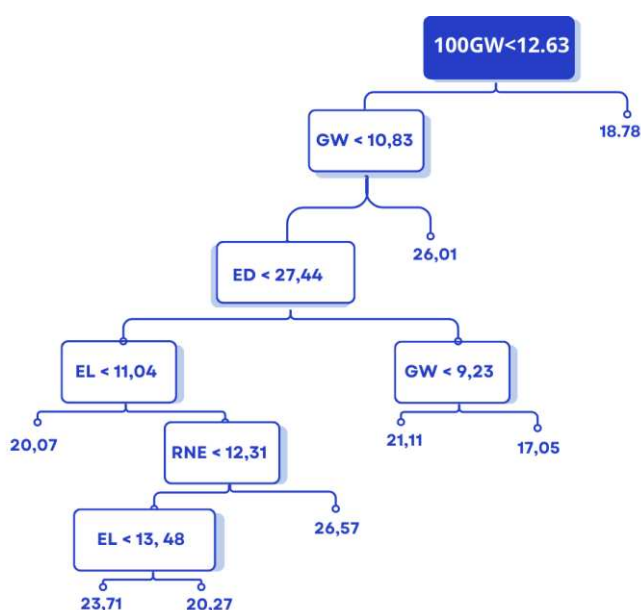


Figura 5 - Árvore de decisão podada para a capacidade de expansão em milho-pipoca, safra de 2021, sob condição hídrica WW. 100GW – Massa de 100 grãos; ED – Diâmetro de espigas; EL – Comprimento de espigas; RNE – Número de fileiras por espigas. Os valores nos nós terminais representam as médias previstas de capacidade de expansão ($g \cdot ml^{-1}$).

Para mesma safra, sob condição de estresse hídrico (WS) em relação a capacidade de expansão (Figura 6), as variáveis explicativas incluíram o comprimento e diâmetro das espigas (EL e ED), número de grãos por fileira (GNR), número de fileiras por espiga (RNE), número total de grãos por espiga (GNE) e peso de 100 grãos (100GW).

Com base na estrutura da árvore o número total de grãos por espiga (GNE) foi fixado na raiz da árvore, posicionando-se como o principal fator de divisão. Isso indica que essa característica exerce a maior influência inicial sobre a variação observada na PE, sendo, portanto, um forte candidato a critério de seleção indireta.

Na sequência, vem o comprimento da espiga (EL) e do diâmetro da espiga (ED) como variáveis de segunda e terceira ordem hierárquica, respectivamente. A interação entre essas variáveis contribuiu para a formação de grupos genotípicos com diferentes padrões de

expansão, sugerindo que aspectos morfológicos da espiga estão associados a níveis mais altos ou mais baixos de desempenho quanto à PE.

O peso de 100 grãos (100GW), embora presente na estrutura do modelo, apareceu apenas em níveis mais baixos da árvore, o que sugere que sua influência é condicionada a determinados intervalos das variáveis superiores.

O modelo apresentou um coeficiente de determinação (R^2) de 0,51, correlação predito-observado de 0,71 e um raiz do erro quadrático médio (RMSE) de 2,77, indicando desempenho satisfatório.

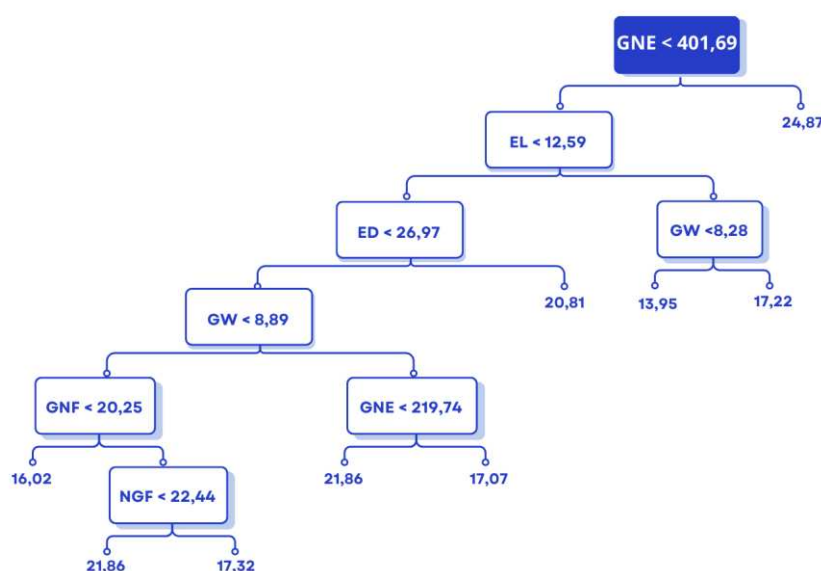


Figura 6 - Árvore de decisão podada para a capacidade de expansão em milho-pipoca, safra de 2021, sob condição hídrica WS. GNF – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; ED – Diâmetro de espigas; EL – Comprimento de espigas; RNE – Número de fileiras por espigas. Os valores nos nós terminais representam as médias previstas de capacidade de expansão ($g \cdot ml^{-1}$).

Embora as árvores de decisão não realizem, de forma explícita, a separação dos efeitos diretos e indiretos entre as variáveis — como é feito na análise de trilha —, sua estrutura em forma de hierarquia permite interpretar, com clareza, quais características secundárias mais influenciam a variável principal. A posição de cada variável nos ramos da árvore e a sequência em que elas aparecem revelam níveis distintos de importância e possíveis interações entre si. Assim, mesmo sem uma decomposição formal, a árvore oferece um caminho interpretativo que, na prática, se assemelha ao raciocínio de causa e efeito adotado em conceitos clássicos.

Ainda que métodos mais sofisticados de aprendizado de máquina sejam recomendados quando o foco está na melhoria do desempenho preditivo — por meio da otimização de métricas como erro quadrático médio e R^2 —, o uso de árvores de decisão com abordagem interpretativa, sem finalidades adicionais, se mostra suficiente e adequado para fins exploratórios na classificação de variáveis. Além disso, sua representação gráfica contribui para a tomada de decisão prática por parte de pesquisadores e produtores, traduzindo relações complexas em caminhos decisórios objetivos e acessíveis (PARSONS et al., 2024). A Figura 7 apresenta, lado a lado, um diagrama de análise de trilha e o perfil de uma árvore de decisão, permitindo ao leitor comparar visualmente as abordagens e compreender de forma mais intuitiva como as relações entre variáveis são representadas em cada método.

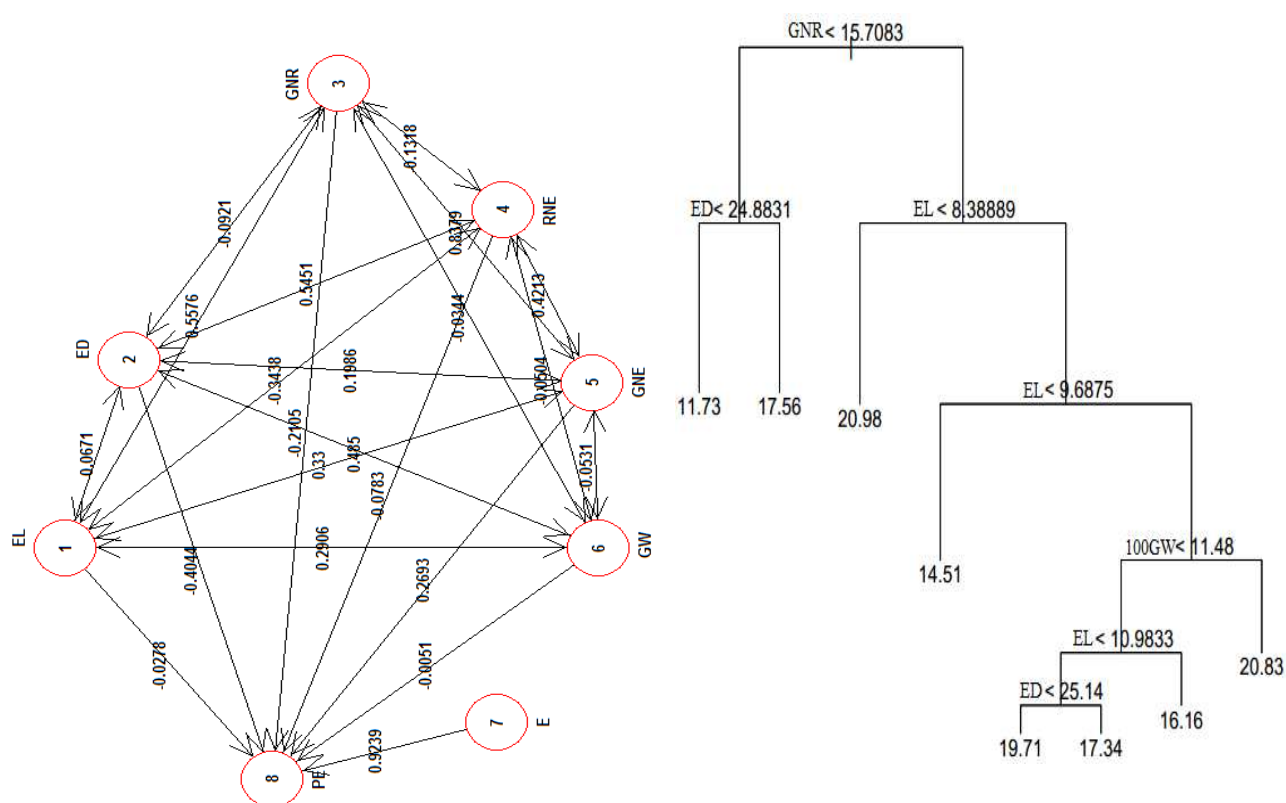


Figura 7. Comparação entre dois métodos de análise utilizados para interpretação de relações entre características auxiliares e variáveis de interesse em milho-pipoca. À esquerda, exemplo de diagrama de análise de trilha, representando relações lineares diretas e indiretas entre características e o caráter principal (setas com coeficientes de efeito). À direita, exemplo de árvore de decisão, ilustrando divisões sucessivas no conjunto de dados e caminhos hierárquicos que levam ao melhor desempenho do caráter principal. A figura destaca as diferenças visuais e interpretativas entre os métodos.

Estudos como o de Shekoofa et al. (2014) já destacaram o potencial das árvores na identificação de características fisiológicas e agrônômicas associadas ao rendimento de grãos, reforçando sua utilidade no mapeamento de variáveis importantes e na definição de critérios de seleção. Assim, as análises apresentadas nesta seção evidenciam que o perfil da árvore pode ser interpretado como uma representação funcional dos caminhos de influência entre características, contribuindo com informações valiosas para programas de melhoramento de milho-pipoca sob diferentes condições ambientais.

Para reforçar essa proposta de complementaridade entre os métodos, foi elaborada uma tabela comparativa entre a análise de trilha e a árvore de decisão, disponível no Apêndice B desta dissertação. Nessa tabela, são apresentados, para cada abordagem, os valores de R^2 obtidos e os melhores caminhos identificados em termos de influência sobre as variáveis de interesse. A comparação evidencia não apenas a eficácia de ambos os métodos na explicação da variabilidade da capacidade de expansão e do rendimento de grãos, mas também ressalta como suas estruturas analíticas distintas — uma baseada em equações e outra em divisões hierárquicas — podem se complementar na interpretação dos dados.

4.4.3. Comparação

A análise de trilha tem sido, ao longo do tempo, uma ferramenta muito utilizada no melhoramento de plantas por oferecer uma visão clara das relações causais entre as variáveis estudadas. Por meio dela, é possível avaliar os efeitos diretos e indiretos de diferentes características auxiliares sobre um caráter principal, sempre partindo do pressuposto de um modelo linear. Isso permite ao pesquisador entender, com base em números, como alterações em uma variável podem impactar diretamente outras, oferecendo argumentos sólidos para decisões estratégicas no processo de seleção (HÜNERMUND et al., 2021). Ainda assim, essa abordagem parte da premissa de que as relações entre as variáveis são lineares, o que pode levar a interpretações errôneas quando, na prática, as interações são mais complexas ou não seguem um padrão linear (GARCIA-CAMPOS et al., 2015).

As árvores de decisão, por outro lado, trazem uma visão diferente e, em muitos casos, mais fortes. Esses modelos trabalham identificando relações não lineares e estruturas hierárquicas, separando o espaço de predição em regiões mais homogêneas. Diferentemente da análise de trilha, elas não se limitam a efeitos somados de maneira simples; conseguem capturar padrões mais sutis e interações implícitas que, muitas vezes, passam despercebidas em análises tradicionais — algo especialmente comum em dados biológicos, onde as

características costumam se relacionar de maneira complexa (DE'ATH; FABRICIUS, 2000; CHATTERJEE, S.; SIMONOFF, J, 2022). Outro ponto a favor é a representação visual em forma de árvore, que torna muito mais fácil compreender como diferentes variáveis contribuem para o resultado.

A seguir são apresentadas tabela comparativa entre a análise de trilha e a árvore de decisão. Na tabela 24 são apresentados, para cada abordagem, os melhores caminhos identificados em termos de influência sobre as variáveis de interesse.

Tabela 24. Importância de característica auxiliares sobre o rendimento de grãos e capacidade de expansão de linhagens de milho pipoca em diferentes cenários

Cenário		Interpretação ww 2020
1	A.Trilha	DC com 0,3765 Efeito Direto 0,3607 FLAV com 0,2052 Efeito direto 0,2404
	Árvore	DC
2	A.Trilha	ANHT com 0,225 Efeito direto 0,3881
	Árvore	AE > NRS
3	A.Trilha	NGF com 0,4056 Efeito direto 0,6971 NGE com 0,3765 Efeito direto 0,3687
	Arvore	NGF > PCG
4	A.Trilha	NGF 0,0474 Efeito direto 0,40
	Árvore	NFE > DESP
		Interpretação ws 2020
1	A.Trilha	SPAD 0,3299 Efeito direto 0,3356 ARC 0,1831 Efeito direto 0,1818
	Árvore	SPAD
2	A.Trilha	Fv/Fm 0,2942 Efeito direto 0,2146 ARC 0,1856 Efeito direto 0,1977
	Árvore	Fv/Fm > NRC > ANTH
3	A.Trilha	NGE 0.3493 Efeito direto 1,7582 DESP 0.2601 Efeito direto 0,0943
	Arvore	NGF > PCG > CESP
4	A.Trilha	PCG 0.1258 Efeito direto 0.1681
	Árvore	NGF > CESP

Interpretação ww 2021		
1	A.Trilha Árvore	ANTH 0.1361 Efeito direto 0.1635 ARC 0.0823 Efeito direto DC > AE > FLAV
2	A.Trilha Árvore	SPAD 0.2139 Efeito direto 0.2033 AP 0.2063 Efeito direto 0.2772 SPAD > ARC > ANTH
3	A.Trilha Árvore	NFE x RG 0.4904 Efeito direto 1.6471 DESP
4	A.Trilha Árvore	CESP x CE 0.066 Efeito direto 0.1173 PCG > DESP > CESP > NFE

Interpretação ws 2021		
1	A.Trilha Árvore	ANTH x RG 0.2342 Efeito direto 0.3178 AP x RG 0.2257 Efeito direto 0.2669 FLAV > ANTH > Fv/Fm > AE > SPAD > ARS > Fv/Fm
2	A.Trilha Árvore	ARC x CE 0.0984 Efeito direto 0.0838 ARS x CE 0.067 Efeito direto 0.03932 SPAD
3	A.Trilha Árvore	NFE x RG 0.3878 Efeito direto 0.7643 DESP x RG 0.2824 Efeito direto 0.07837 NFE > PCG > CESP > NGE > NFE
4	A.Trilha Árvore	PCG x CE 0.3883 Efeito direto 0.3341 DESP x CE 0.1588 Efeito direto 0.0131 NGE

- 1: Características Morfológicas e Rendimento de Grão
 2: Características Morfológicas e Capacidade de Expansão
 3: Características Agronômicas e Rendimento de Grão
 4: Características Agronômicas e Capacidade de Expansão

Os resultados deixam essa diferença bem evidente. A análise de trilha destacou variáveis importantes ao mostrar seus efeitos diretos e indiretos dentro de um contexto linear. Já as árvores de decisão revelaram caminhos diferentes, muitas vezes organizados de forma hierárquica, em que novas variáveis surgem como fundamentais ao lado das já conhecidas. Em alguns casos, as duas metodologias apontaram para a mesma característica principal; em

outros, a árvore trouxe à tona variáveis que não se destacaram na análise de trilha, sugerindo a presença de interações ou dependências mais complexas que o modelo linear não foi capaz de captar.

Por fim, destaca-se que, em culturas como o milho-pipoca, em que o desafio é melhorar simultaneamente o rendimento e a qualidade em relação ao estresse, a combinação de diferentes ferramentas estatísticas e computacionais se mostra não só desejável, mas necessária. A complementaridade entre a análise de trilha e abordagens modernas pode ser a chave para decisões mais assertivas e seleções mais eficientes em programas de melhoramento genético. Diante dos resultados obtidos com a análise de trilha, observou-se que sua aplicação se mostrou eficaz na decomposição dos efeitos totais em componentes diretos e indiretos, possibilitando a identificação de características secundárias com efeito causal sobre os caracteres principais. No entanto, considerando as limitações inerentes ao modelo linear adotado por essa abordagem, sobretudo em contextos com interações complexas entre variáveis, esse trabalho propôs o uso complementar das árvores de decisão como uma alternativa metodológica ao tradicional diagrama de causa e efeito. A topologia das árvores, por sua estrutura hierárquica, permite não apenas estimar o grau de relacionamento entre variáveis principais (como rendimento de grãos e capacidade de expansão) e características auxiliares, mas também visualizar interações condicionais e a importância relativa de cada variável na modelagem preditiva.

5. CONCLUSÕES

- Há variabilidade genética para todas as características mensuradas demonstrando potencialidade dos genótipos para fins de melhoramento e para inferência sobre a relevância de características auxiliares para seleção do rendimento de grãos e da capacidade de expansão.
- O posicionamento do desempenho de genótipos foi influenciado pela condição de fornecimento de água no experimento. Genótipos destacado em grupos elite em condições favoráveis demonstraram desempenho inferior em condições de estresse hídrico.
- Apesar de baixa frequência foi possível identificar genótipos de bom desempenho quanto a rendimento de grãos e capacidade de expansão em condições adequadas de fornecimento de água ou em déficit hídrico.
- A importância de características auxiliares é influenciada pelas condições ambientais.
- A análise fundamentada em árvore de decisão permitiu reconhecer a maior relevância dos grupos de variáveis auxiliares sobre o rendimento e capacidade de expansão do milho pipoca.
- As topologias da árvore de decisão mostraram ser de grande utilidade para o entendimento da atuação de características auxiliares sobre características de interesse.

A árvore de decisão foi uma metodologia superior à análise de trilha para indicar características auxiliares seleção indireta do rendimento de grão e de capacidade de expansão de milho-pipoca.

REFERÊNCIAS

- ANSARIFAR, J.; WANG, L.; ARCHONTOULIS, S. V. **Um modelo de regressão de interação para previsão de produtividade agrícola.** *Scientific Reports*, v. 11, 17754, 2021. DOI: <https://doi.org/10.1038/s41598-021-97221-7>.
- BANERJEE, B. P. et al. **Machine learning regression analysis for estimation of crop emergence using multispectral UAV imagery.** *Remote Sensing*, v. 13, n. 15, p. 2918, 2021. DOI: <https://doi.org/10.3390/rs13152918>.
- BASU, B. J. et al. **Correlation, regression and path coefficient analysis between morphological characteristics and grain yield of popcorn (*Zea mays everta*) under different levels of fertilizer and plant density.** *Journal of Pharmacognosy and Phytochemistry*, v. 6, p. 662–664, 2017. Disponível em: <http://www.phytojournal.com/archives/2017/vol6issue6S/PartP/SP-6-6-147-545.pdf>. Acesso em: 6 maio 2025.
- BISPO, R. B. **Déficit hídrico na estimação da capacidade combinatória de linhagens de milho-pipoca para características agrônômicas e radiculares.** 2020. Dissertação (Mestrado) – Universidade Federal de Viçosa, Viçosa, MG, 2020.
- BOKER, S. M.; McARDLE, J. J. **Path analysis and path diagrams.** *Wiley StatsRef: Statistics Reference Online*, 2014. DOI: <https://doi.org/10.1002/9781118445112.stat06517>.
- CABRAL, P. D. S. et al. **Cause and effect of quantitative characteristics on grain expansion capacity in popcorn.** *figshare*, 2021. DOI: <https://doi.org/10.6084/m9.figshare.14326912>.
- CARANHATO, A. L. H. et al. **Genetic improvement in popcorn.** *Revista Brasileira de Milho e Sorgo*, v. 21, e1272, 2022. DOI: <https://doi.org/10.18512/1980-6477/rbms.v21.e1272>.
- CARVALHO, C. M. et al. **Seleção precoce para tolerância à seca em milho-pipoca com base em efeitos gênicos estimados em mudas.** *Frontiers in Plant Science*, v. 14, 2023. DOI: <https://doi.org/10.3389/fpls.2023.1203972>.
- CHATTERJEE, S.; SIMONOFF, J. **Tree-based models.** *Institution of Engineering and Technology eBooks*, 2022. p. 105–128. DOI: https://doi.org/10.1049/pbtr038e_ch9.
- CHAUHAN, N.; SULOCHANA, V.; KUMAR, R. P. **FARMER: Forecasting Agricultural Results with Machine Learning Evaluation and Regression.** 2024. Disponível em: <https://doi.org/10.1109/icrito61523.2024.10522442>. Acesso em: 6 maio 2025.
- COSTA, T. S. **Características nutricional, bioativa, física e sensorial de diferentes genótipos de milho-pipoca.** 2022. Tese (Doutorado) – Universidade Federal de Viçosa, Viçosa, MG, 2022.
- CRUZ, C. D. **GENES – a software package for analysis in experimental statistics and quantitative genetics.** *Acta Scientiarum*, v. 35, n. 3, p. 271–276, 2013.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2012. p. 181–209.

DALLACORT, R. et al. **Distribuição das chuvas no município de Tangará da Serra, médio norte do Estado de Mato Grosso, Brasil**. *Acta Scientiarum Agronomy*, v. 33, n. 2, p. 193–200, 2011. DOI: <https://doi.org/10.4025/actasciagron.v33i2.5838>.

DE OLIVEIRA, U. A. et al. **Revelando linhas de pipoca latino-americanas resistentes à seca por meio de avaliação agrônômica e fisiológica**. *Life*, v. 14, n. 6, p. 743, 2024. DOI: <https://doi.org/10.3390/life14060743>.

DE'ATH, G.; FABRICIUS, K. **Classification and regression trees: A powerful yet simple technique for ecological data analysis**. *Ecology*, s.d. DOI: <https://doi.org/10.1890/0012-965>

DURÃES, F. O. M. et al. **Fenotipagem associada à tolerância a seca em milho**. Sete Lagoas: Embrapa Milho e Sorgo, 2004.

G. et al. **The evolution of path analysis in genetics and behavior**. *Journal of Quantitative Genetics*, v. 45, p. 123–135, 2024.

GAMA, E. E. G. et al. **Milho pipoca**. Brasília: EMBRAPA/CNPMS, 1990.

GAO, Y.; LYNCH, J. P. **Reduced crown root number improves water acquisition under water deficit stress in maize (*Zea mays* L.)**. *Journal of Experimental Botany*, v. 67, p. 4545–4557, 2016.

GARCIA-CAMPOS, M. A.; ESPINAL-ENRÍQUEZ, J.; HERNÁNDEZ-LEMUS, E. **Pathway analysis: State of the art**. *Frontiers in Physiology*, v. 6, p. 383, 2015. DOI: <https://doi.org/10.3389/fphys.2015.00383>.

HÜNERMUND, P.; KAMINSKI, J.; SCHMITT, C. **Causal machine learning and business decision making**. *Social Science Research Network*, 2021. DOI: <https://doi.org/10.2139/SSRN.3867326>.

IZENMAN, A. J. **Recursive partitioning and tree-based methods**. In: _____. *Modern multivariate statistical techniques*. New York: Springer, 2013. p. 281–314. DOI: https://doi.org/10.1007/978-0-387-78189-1_9.

KAMPHORST, S. H. et al. **Comparison of selection traits for effective popcorn (*Zea mays* L. var. everta) breeding under water limiting conditions**. *Frontiers in Plant Science*, v. 11, p. 1289, 2020. DOI: <https://doi.org/10.3389/FPLS.2020.01289>.

KAMPHORST, S. H. et al. **Supporting physiological trait for indirect selection for grain yield in drought-stressed popcorn**. *Plants*, v. 10, n. 8, p. 1510, 2021. DOI: <https://doi.org/10.3390/plants10081510>.

KAMPHORST, S. H. et al. **Water stress adaptation of popcorn roots and association with agronomic traits**. *Genetics and Molecular Research*, v. 17, n. 3, 2018. DOI: <https://doi.org/10.4238/gmr18078>.

KLEM, L. **Análise de caminho**. In: GRIMM, L. G.; YARNOLD, P. R. (Eds.). *Leitura e compreensão de estatísticas multivariadas*. Washington: APA, 1995. p. 65–97.

LEITE, J. T. **Importância relativa de características morfofisiológicas para a seleção de linhagens de milho-pipoca em condição de seca**. 2023. 93 f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, 2023.

LESK, C.; ROWHANI, P.; RAMANKUTTY, N. **Influência de desastres climáticos extremos na produção global de culturas**. *Nature*, v. 529, n. 7584, p. 84–87, 2016. DOI: <https://doi.org/10.1038/nature16467>.

LI, B. et al. **Flavonoides melhoram a tolerância à seca de mudas de milho regulando a homeostase de espécies reativas de oxigênio**. *Plant and Soil*, v. 461, p. 389–405, 2021. DOI: <https://doi.org/10.1007/s11104-020-04814-8>.

LI, X. et al. **Genotypic variability in root morphology in a diverse wheat genotypes under drought and low phosphorus stress**. *Plants*, v. 13, p. 3361, 2024. DOI: <https://doi.org/10.3390/plants13233361>.

LOPEZ-CRUZ, M. et al. **Leveraging data from the Genomes to Fields Initiative to investigate genotype-by-environment interactions in maize in North America**. *Nature Communications*, v. 14, n. 1, p. 6904, 2023. DOI: 10.1038/s41467-023-42687-4.

MACHADO, J. R. A.; FONTANELI, R. S. **Inserção das culturas de milho e sorgo na agricultura familiar na região sul brasileira**. 2014.

MACHADO, M. F. P. S. **Biochemical composition of the pericarp cell wall of popcorn inbred lines with different popping expansion**. *Current Research in Food Science*, v. 5, p. 102–106, 2022. DOI: <https://doi.org/10.1016/j.crfs.2021.12.011>.

MASTRANGELO, A. M. et al. **Improvement of drought resistance in crops: from conventional breeding to genomic selection**. Springer, Dordrecht, 2012. DOI: https://doi.org/10.1007/978-94-007-2220-0_7.

MENEGALDO, J. G. **A importância do milho na vida das pessoas**. *Agrosoft Brasil*, 2011. Disponível em: <http://www.agrosoft.org.br/agropag/217861.htm>. Acesso em: 30 abr. 2025.

MONTESINOS-LÓPEZ, O. A. et al. **Application of machine learning algorithms in plant breeding: a review**. *Frontiers in Plant Science*, v. 11, 2020. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpls.2020.624273/full>. Acesso em: 6 maio 2025.

MORETTIN, P. A.; SINGER, J. M. **Estatística e ciência de dados**. Rio de Janeiro: LTC, 2022. p. 545–592.

MULAIK, S. A. **History of path analysis**. *Wiley StatsRef: Statistics Reference Online*, 2014. DOI: <https://doi.org/10.1002/9781118445112.stat06308>.

PACHECO, C. A. P.; CASTOLDI, F. L.; ALVARENGA, E. M. **Efeito do dano mecânico na qualidade fisiológica e na capacidade de expansão de sementes de milho pipoca.** Viçosa: UFV, 1996.

PAINEL INTERGOVERNAMENTAL SOBRE MUDANÇAS CLIMÁTICAS – IPCC. **Mudanças climáticas 2022: impactos, adaptação e vulnerabilidade.** Cambridge: Cambridge University Press, 2022. DOI: <https://doi.org/10.1017/9781009325844>.

PATH ANALYSIS. **Path analysis.** In: *Multivariate Statistics for the Social Sciences.* Routledge eBooks, 2022. p. 159–176. DOI: <https://doi.org/10.4324/9781003343097-7>.

PARSONS, J.; HEWLETT, J.; TRANEL, J. **Branching out: Harnessing the power of decision trees.** 2024. DOI: <https://doi.org/10.32873/unl.dc.cap044>.

PHALAK, M. P.; BHANDARI, M. K.; SHARMA, R. **Analysis of decision tree—a survey.** *International Journal of Engineering Research and Technology*, v. 3, n. 3, 2014. Disponível em: <https://www.ijert.org/research/analysis-of-decision-tree-a-survey-IJERTV3IS030302.pdf>. Acesso em: 6 maio 2025.

PIYAR, S. N.; BAULKANI, S. **Crop prediction based on soil and environmental features using classification models.** 2024. DOI: <https://doi.org/10.1109/icdsis61070.2024.10594482>.

RANGEL, R. M.; AMARAL JÚNIOR, A. T. D.; FREITAS JÚNIOR, S. D. P. **Associação entre características agronômicas e capacidade de expansão em população de milho pipoca sob seleção recorrente.** *Ciência e Agrotecnologia*, v. 35, p. 225–233, 2011.

RIBEIRO, T. A. et al. **Estudos de arquitetura radicular em milho-pipoca sob déficit hídrico.** *Revista Ciência Agronômica*, v. 55, 2024.

RINALDI, P. C. N. et al. **Trait prediction through computational intelligence and machine learning applied to the improvement of white oat (*Avena sativa* L.).** *Revista Ceres*, v. 69, n. 2, p. 229–238, 2022. Disponível em: <https://www.scielo.br/j/rceres/a/qkF7VXhK6VvqFvk9DSS8HLc>. Acesso em: 6 maio 2025.

SANTOS, T. O. et al. **Phenotyping Latin American open-pollinated varieties of popcorn for environments with low water availability.** *Plants*, v. 10, n. 6, p. 1211, 2021. DOI: <https://doi.org/10.3390/PLANTS10061211>.

SAKIN, M. A. et al. **Efeitos do tipo de cultivar na produtividade e qualidade do milho-pipoca (*Zea mays* everta).** *Revista de Ciências Hortícolas e de Cultivos da Nova Zelândia*, v. 33, n. 1, p. 17–23, 2005. DOI: <https://doi.org/10.1080/01140671.2005.9514326>.

SAWAZAKI, E. **Parâmetros genéticos em milho pipoca (*Zea mays* L.).** 1996. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 1996.

SHEKOOFA, A. et al. **Determining the most important physiological and agronomic traits contributing to maize grain yield through machine learning algorithms: a new avenue in intelligent agriculture.** *PLOS ONE*, v. 9, n. 5, 2014. DOI: <https://doi.org/10.1371/journal.pone.0097288>.

SONG, A. et al. **Efeitos do tamanho do grão e do genótipo no volume de estouro da pipoca e no número de grãos não estourados.** *Cereal Chemistry*, v. 68, n. 5, p. 464–467, 1991.

STREINER, D. L. **Finding our way: an introduction to path analysis.** *The Canadian Journal of Psychiatry*, v. 50, n. 2, p. 115–122, 2005. DOI: <https://doi.org/10.1177/070674370505000207>.

TINOVONGA G. et al. **Avaliação da diversidade genética e alinhamento heterótico de linhas endogâmicas de milho CIMMYT e IITA adaptadas à África Subsaariana.** *Crop Science*, v. 65, n. 1, 2024. DOI: <https://doi.org/10.1002/csc2.21401>.

THOMAS, T.; VIJAYARAGHAVAN, A. P.; EMMANUEL, S. **Applications of decision trees.** In: *Decision Trees*. Singapore: Springer, 2020. p. 157–184. DOI: https://doi.org/10.1007/978-981-15-1706-8_9.

VIJAYABHARATHI, A.; ANANDAKUMAR, C. R.; GNANAMALAR, R. P. **Analysis of correlations and path effects for popping expansion in popcorn (*Zea mays* var. *everta* Sturt.).** *Electronic Journal of Plant Breeding*, v. 1, n. 1, p. 60–64, 2009. Disponível em: <https://www.cabdirect.org/cabdirect/abstract/20103117227>. Acesso em: 6 maio 2025.

WASHBURN, J. D. et al. **Machine learning approach for prescriptive plant breeding.** *Scientific Reports*, v. 11, 2021. Disponível em: <https://www.nature.com/articles/s41598-019-53451-4>. Acesso em: 6 maio 2025.

ZHOU, H. **Decision trees.** In: *Machine Learning with Decision Trees*. Apress, 2023. p. 161–185. DOI: https://doi.org/10.1007/978-1-4842-9771-1_10.

APÊNDICE A - CORRELAÇÕES GENOTÍPICAS E AMBIENTAIS

Correlações genotípicas e ambientais das safras 2020 e 2021 em duas condições hídricas (WW e WS)

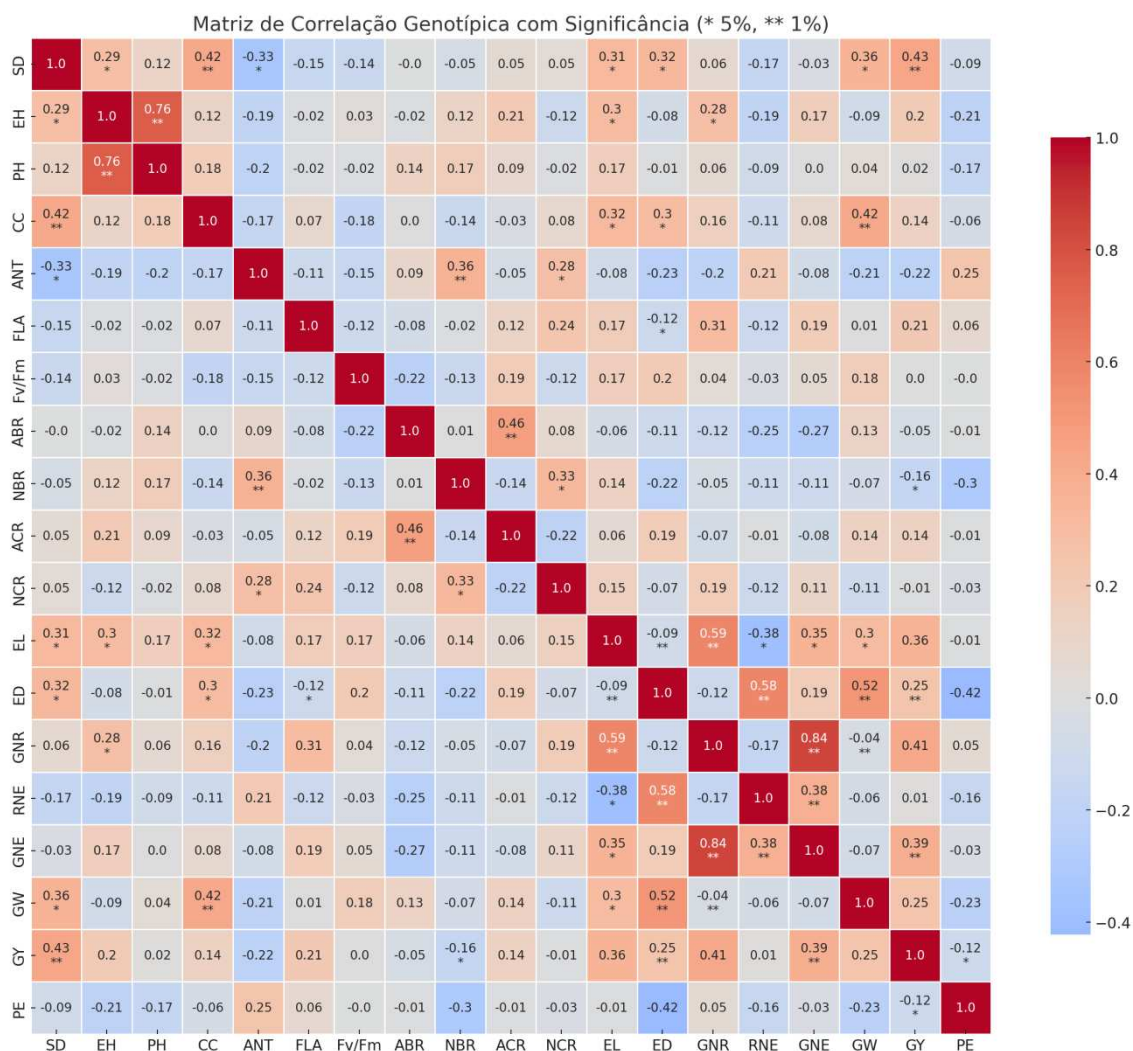


Figura 8. Estimativas de correlações genotípicas entre as características estudadas na condição bem irrigada (WW), na safra de 2020. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.



Figura 9. Estimativas de correlações Ambientais entre as características estudadas na condição bem irrigada (WW), na safra de 2020. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

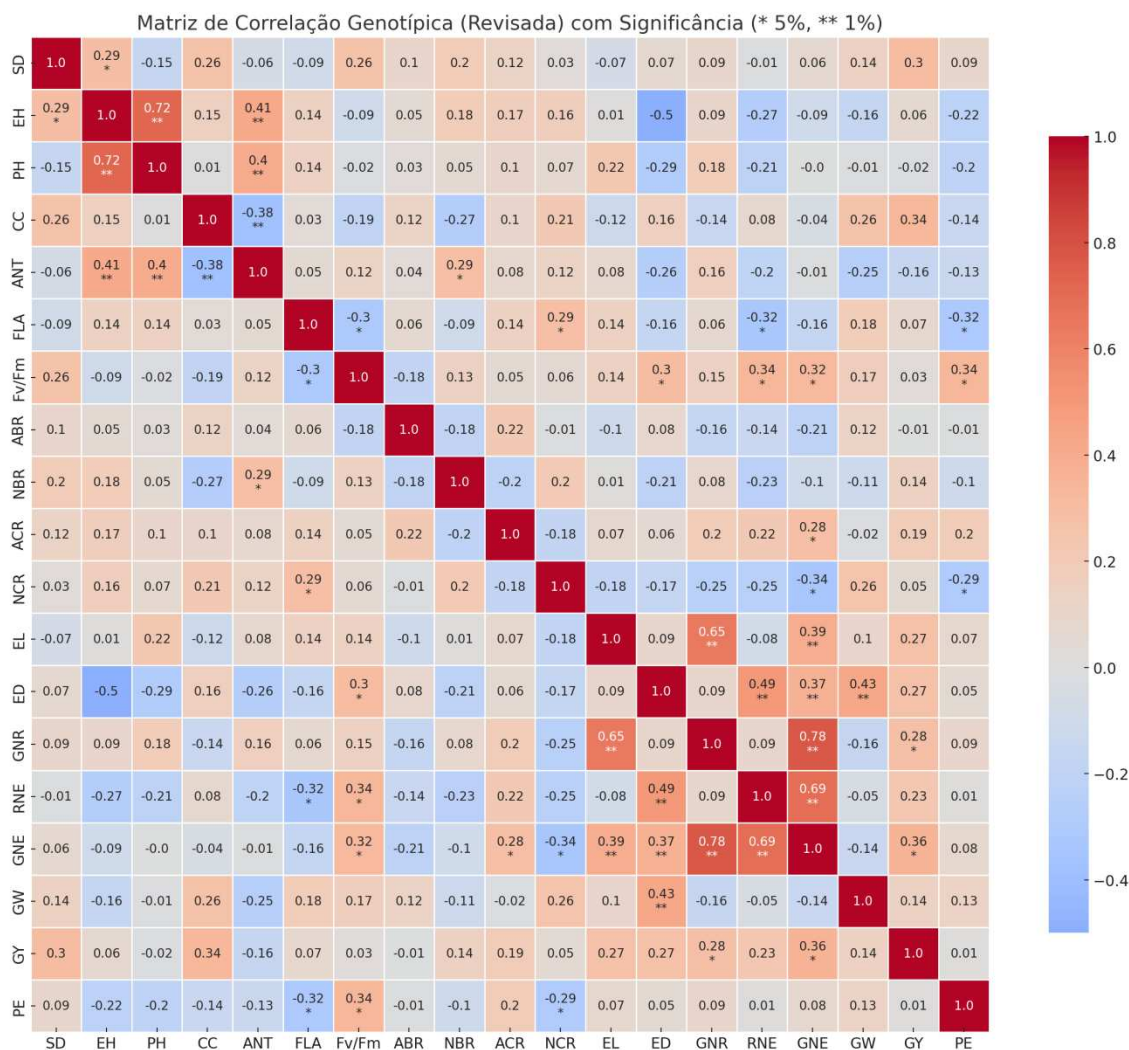


Figura 10. Estimativas de correlações genotípicas entre as características estudadas na condição com déficit hídrico (WS), na safra de 2020. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

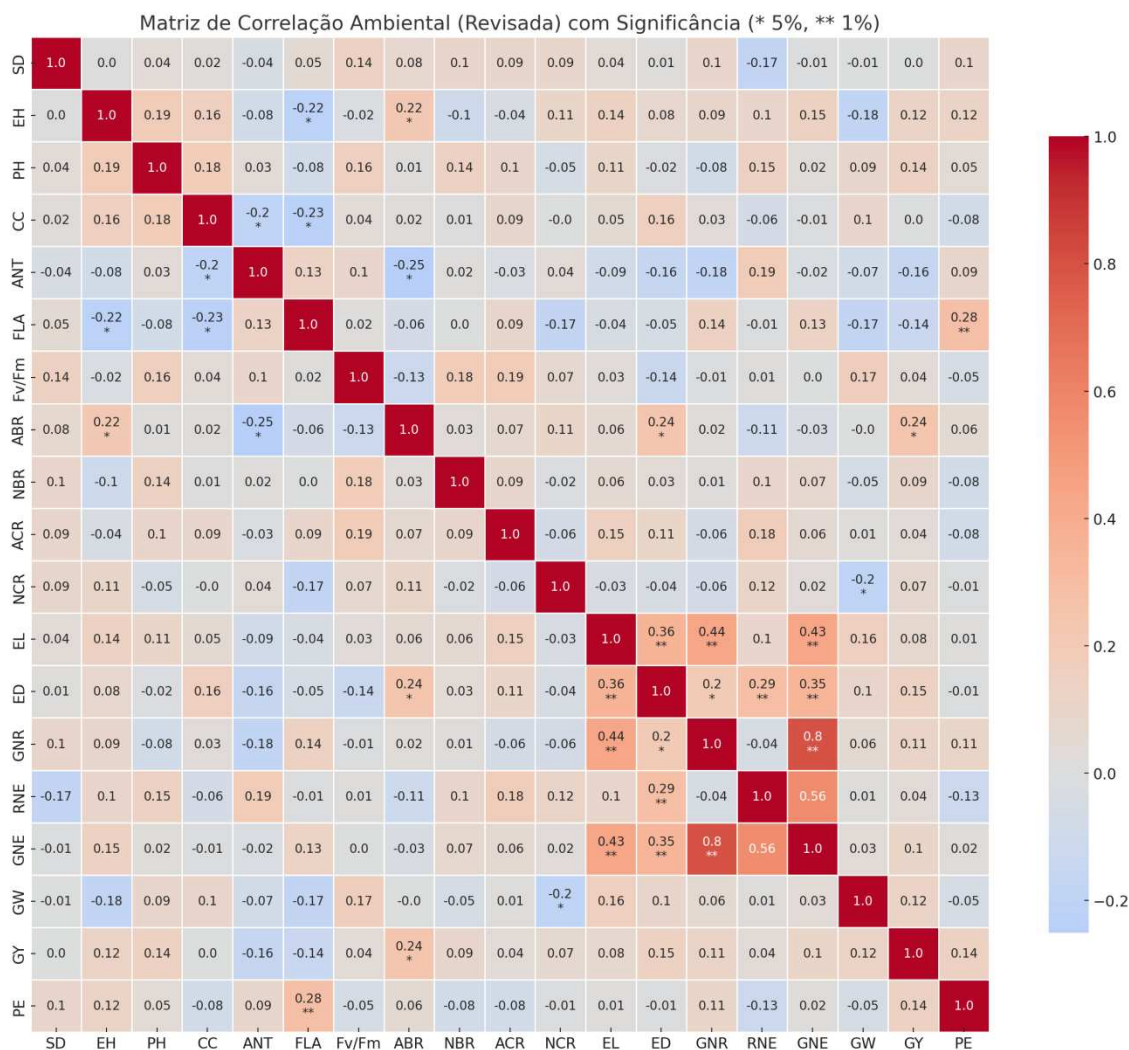


Figura 11. Estimativas de correlações ambiental entre as características estudadas na condição com déficit hídrico (WS), na safra de 2020. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

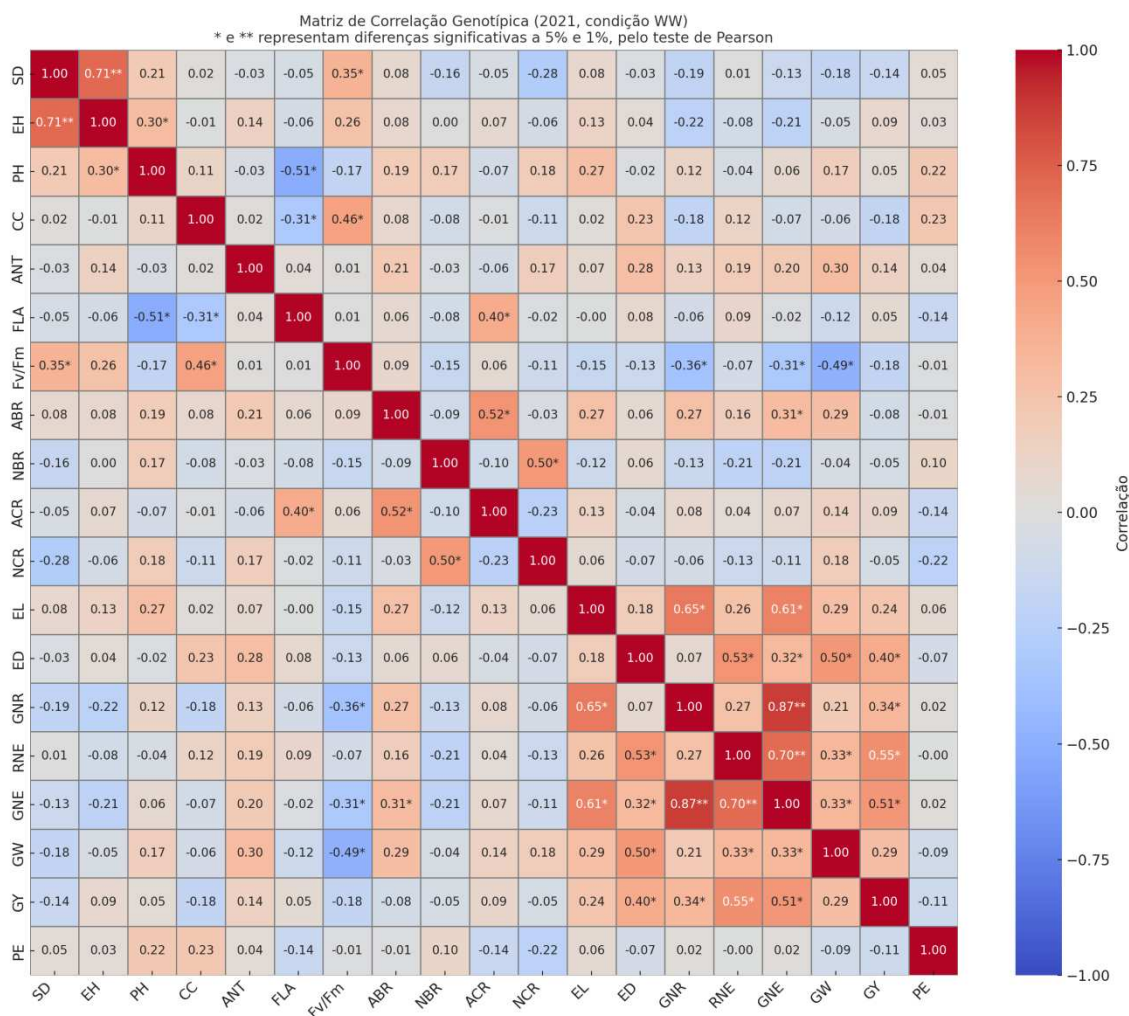


Figura 12. Estimativas de correlações genotípica entre as características estudadas na condição bem irrigada (WW), na safra de 2021. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

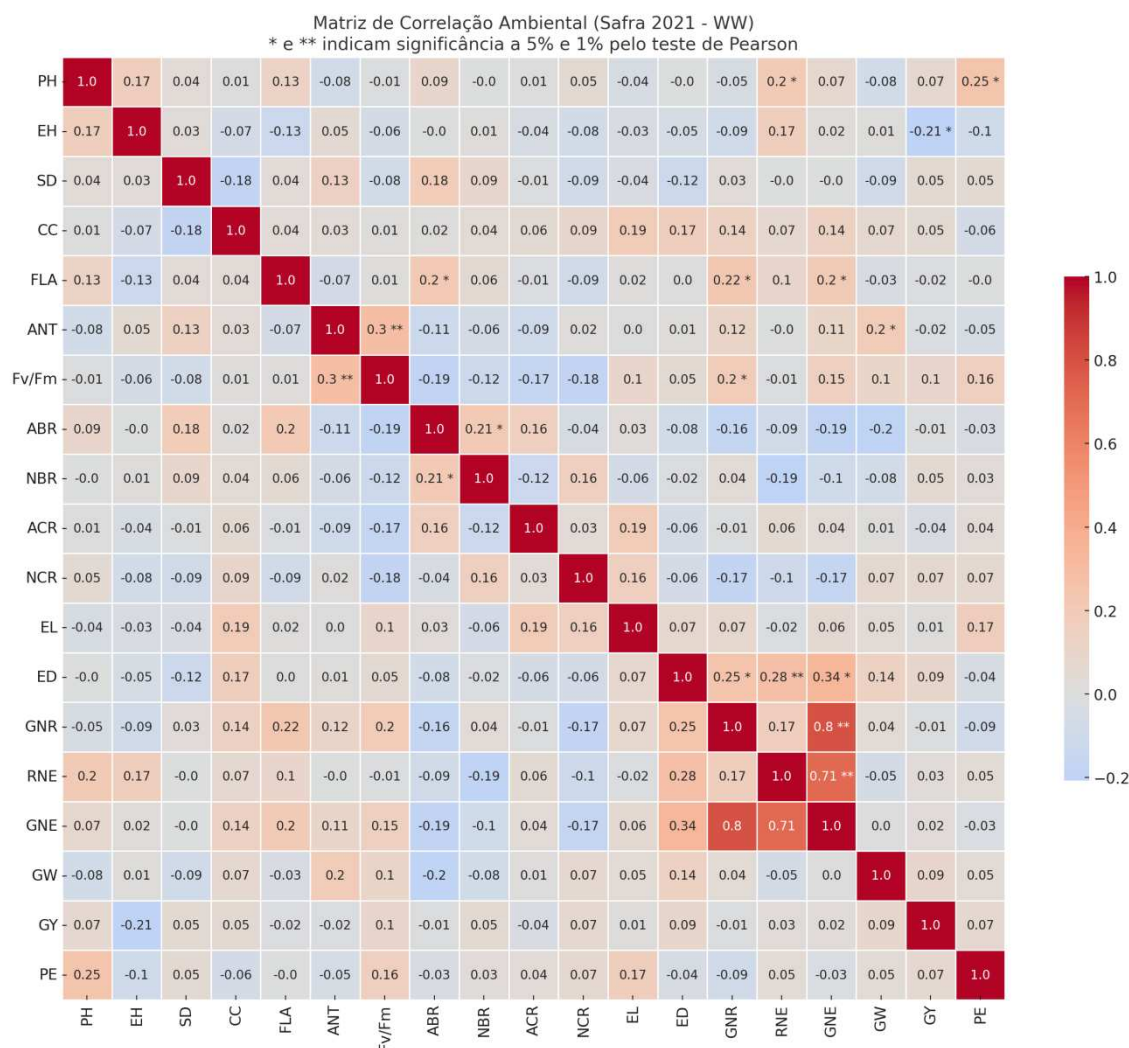


Figura 13. Estimativas de correlações ambiental entre as características estudadas na condição bem irrigada (WW), na safra de 2021. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

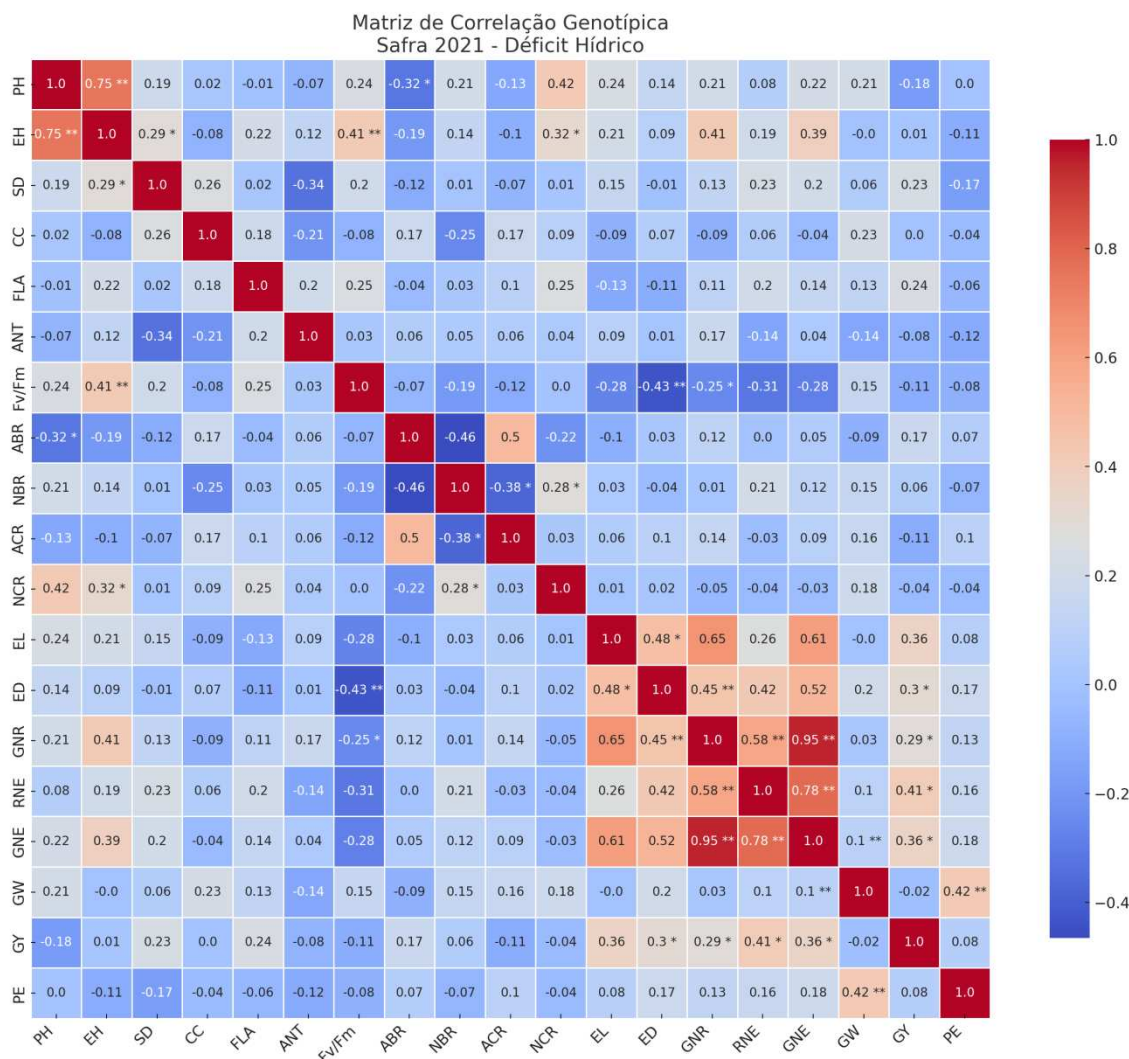


Figura 14. Estimativas de correlações genotípicas entre as características estudadas na condição bem irrigada (WS), na safra de 2021. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espigas; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

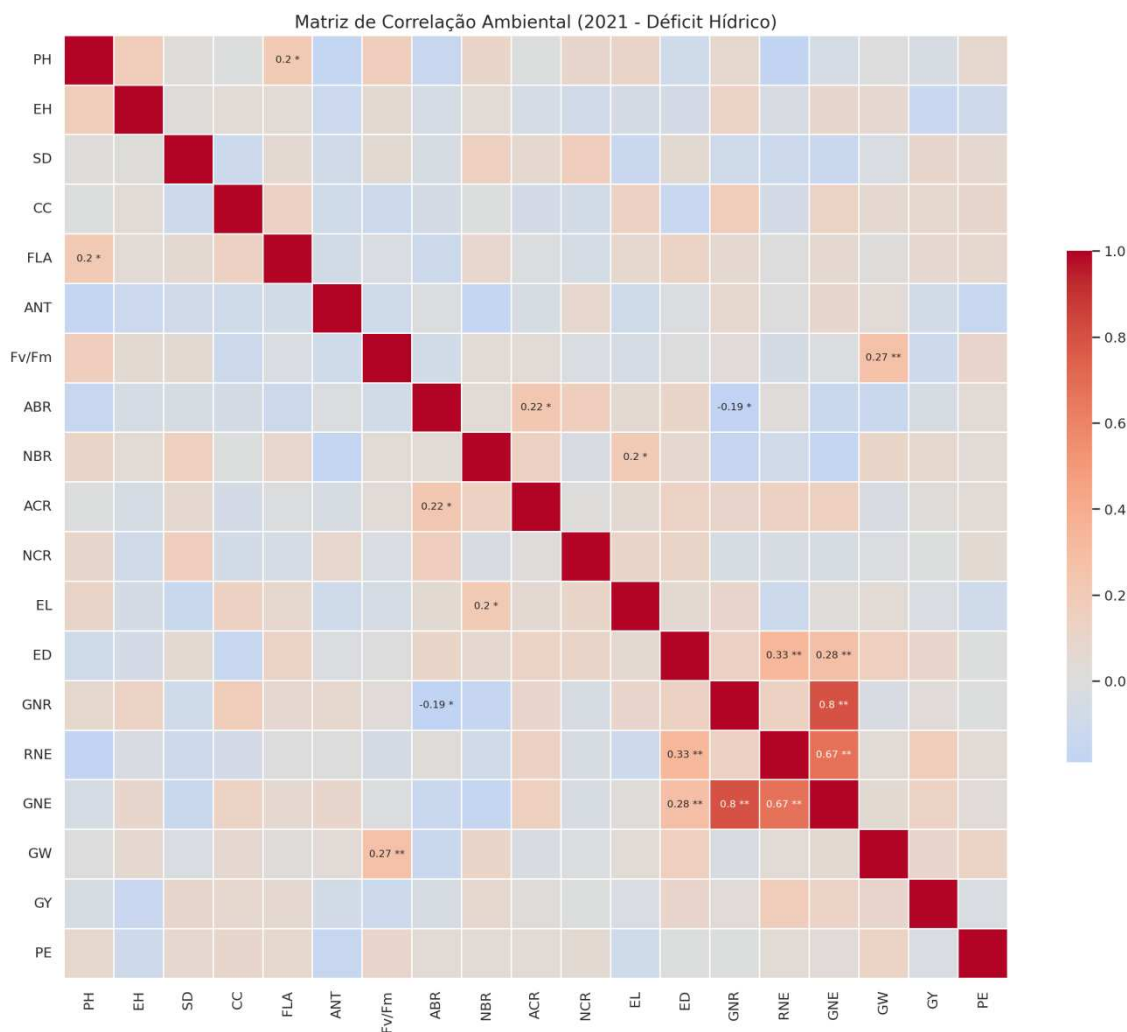


Figura 15. Estimativas de correlações Ambientais entre as características estudadas na condição bem irrigada (WS), na safra de 2021. EH – Altura de espigas; PH – Altura de plantas; SD – Diâmetro de colmo; EL – Comprimento de espigas; ED – Diâmetro de espigas; RNE – Número de fileiras por espiga; GNR – Número de grãos por fileira; GNE – Número de grãos por espiga; 100GW – Massa de 100 grãos; GY – Rendimento de grãos; PE – Capacidade de expansão; CC – Teor relativo de clorofilas; ANT – Teor relativo de antocianinas; FLA – Teor relativo de flavonoides; Fv/Fm – Fluorescência de clorofilas; ABR – Ângulo de raízes de suporte; ACR – Ângulo de raízes da coroa; NBR – Número de raízes de suporte; NCR – Número de raízes da coroa. * e **, representam diferenças significativas em nível de 5 e 1%, pelo teste t.

APÊNDICE B – DIAGRAMAS DE ÁRVORE DE DECISÃO

Diagramas da Árvore de Decisão nas Safras de 2020 e 2021 sob Condições Contrastantes de Disponibilidade Hídrica.

Safra 2020, na condição WW.

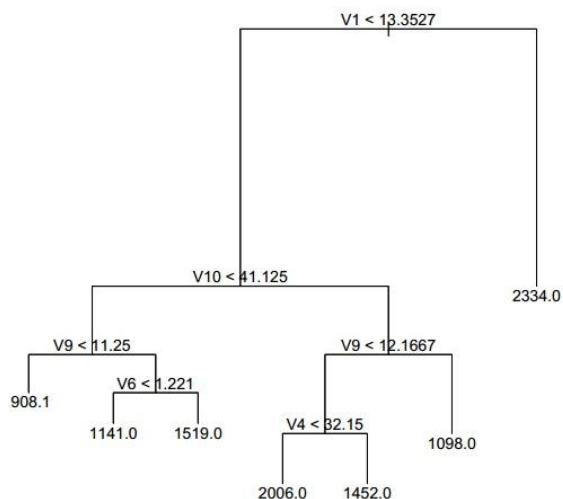


Figura 16. Árvore de decisão para o rendimento de grãos (RG) e variáveis morfológicas em milho-pipoca, safra de 2020, sob condição hídrica WW.

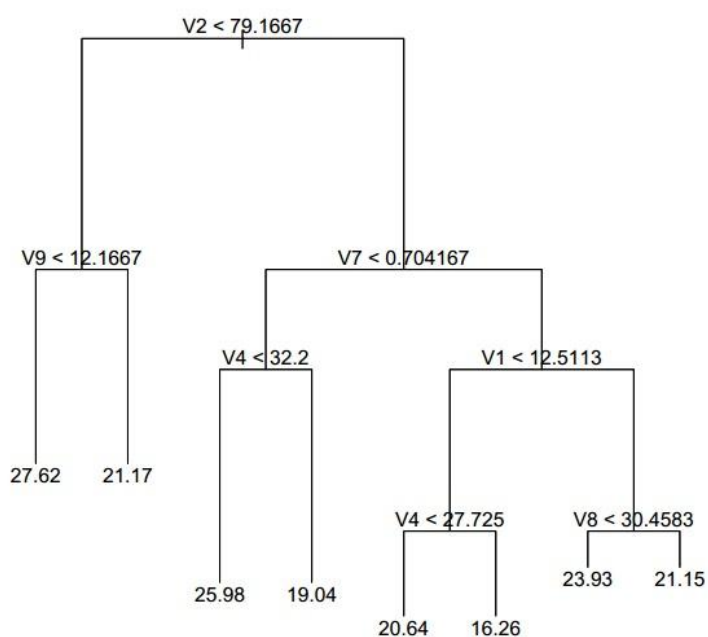


Figura 17. Árvore de decisão para capacidade de expansão e variáveis morfológicas em milho-pipoca, safra de 2020, sob condição hídrica WW.

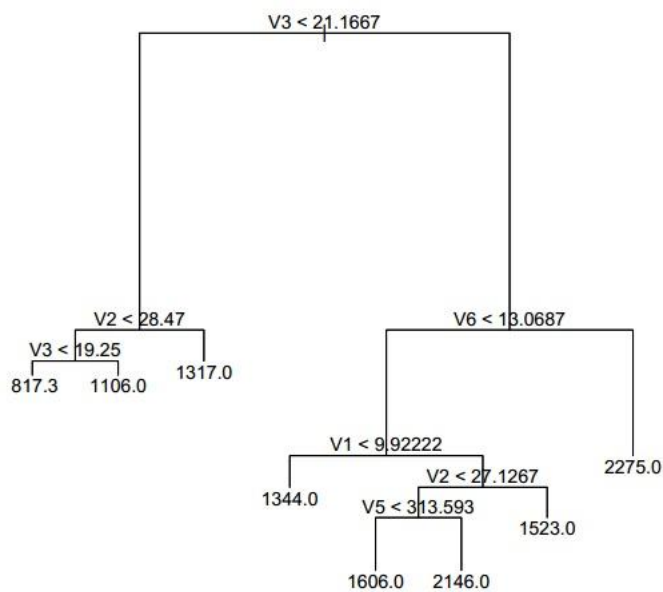


Figura 18. Árvore de decisão para GY e variáveis agrônômicas em milho-pipoca, safra de 2020, sob condição hídrica WW.

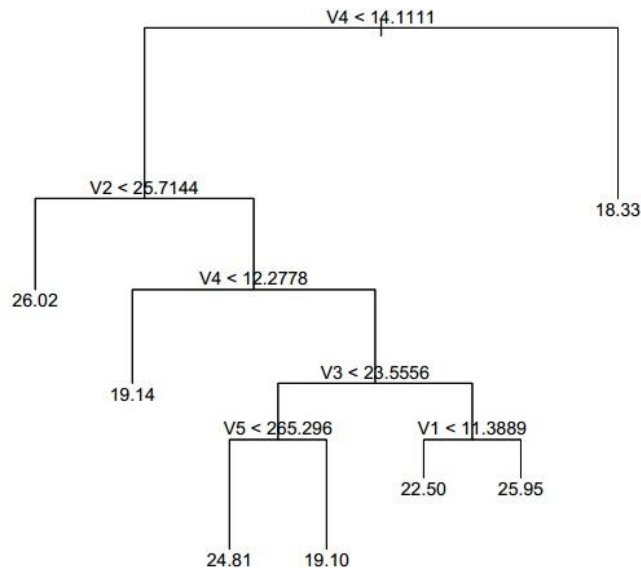


Figura 19. Árvore de decisão para capacidade de expansão e variáveis agrônômicas em milho-pipoca, safra de 2020, sob condição hídrica WW.

Safra 2020, na condição WS.

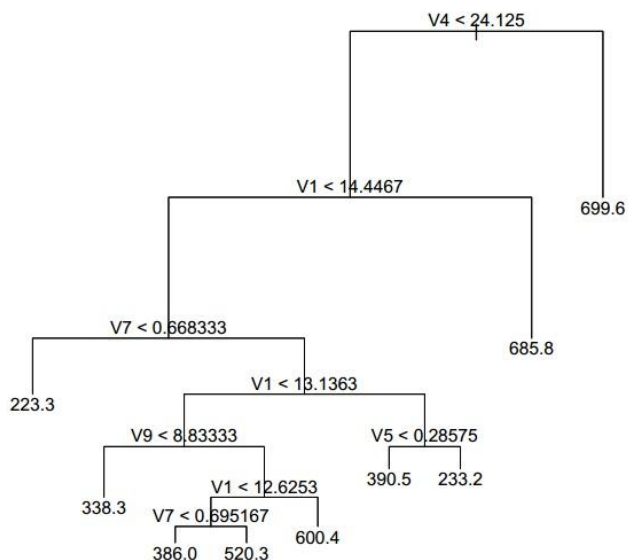


Figura 20. Árvore de decisão para GY e variáveis morfológicas em milho-pipoca, safra de 2020, sob condição hídrica WS.

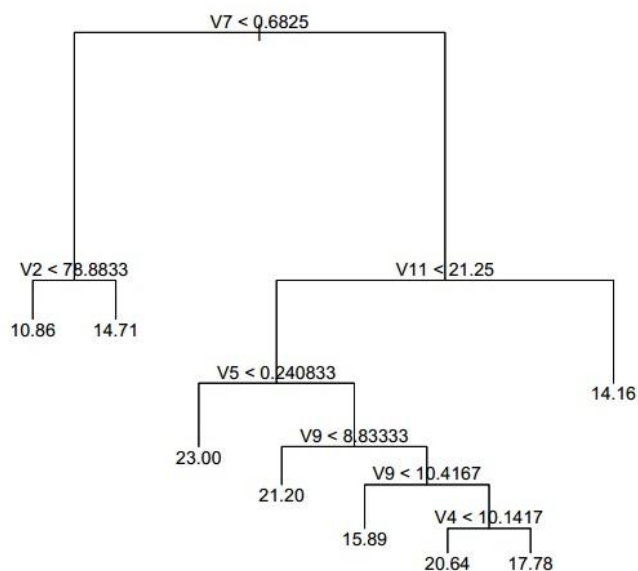


Figura 21. Árvore de decisão para capacidade de expansão e variáveis morfológicas em milho-pipoca, safra de 2020, sob condição hídrica WS.

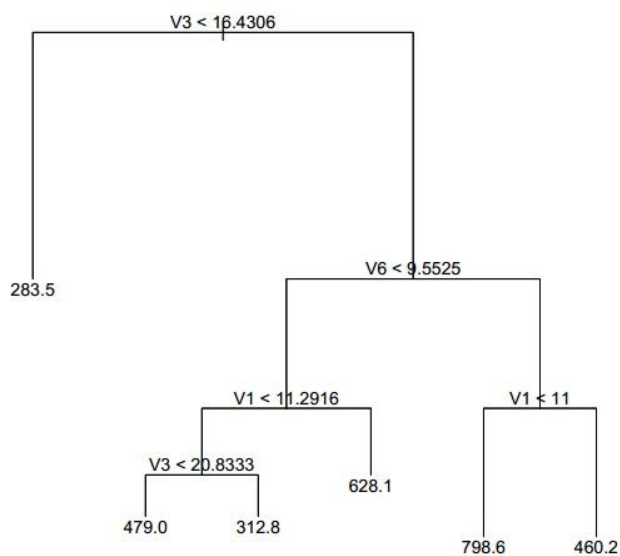


Figura 22. Árvore de decisão para GY e variáveis agrônômicas em milho-pipoca, safra de 2020, sob condição hídrica WS.

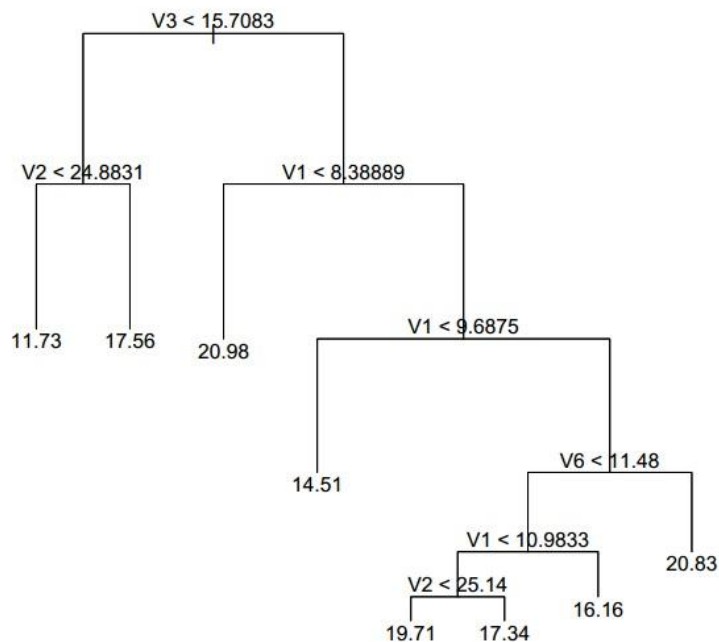


Figura 23. Árvore de decisão para PE e variáveis agrônômicas em milho-pipoca, safra de 2020, sob condição hídrica WS.

Safra 2021, na condição WW.

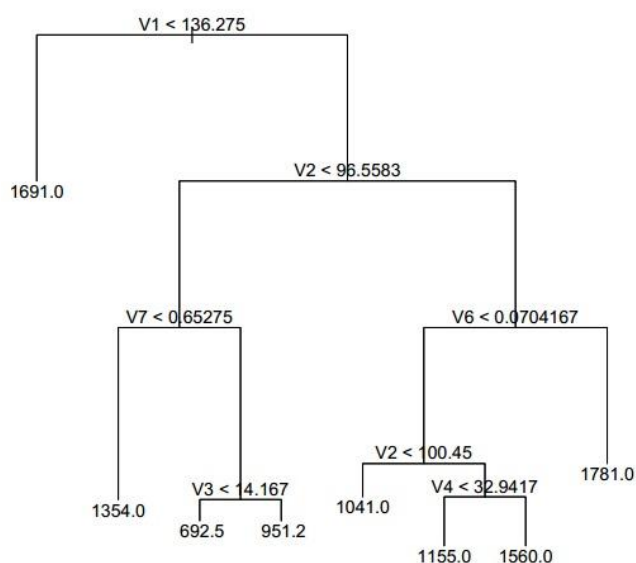


Figura 24. Árvore de decisão para GY e variáveis morfológicas em milho-pipoca, safra de 2021, sob condição hídrica WW.

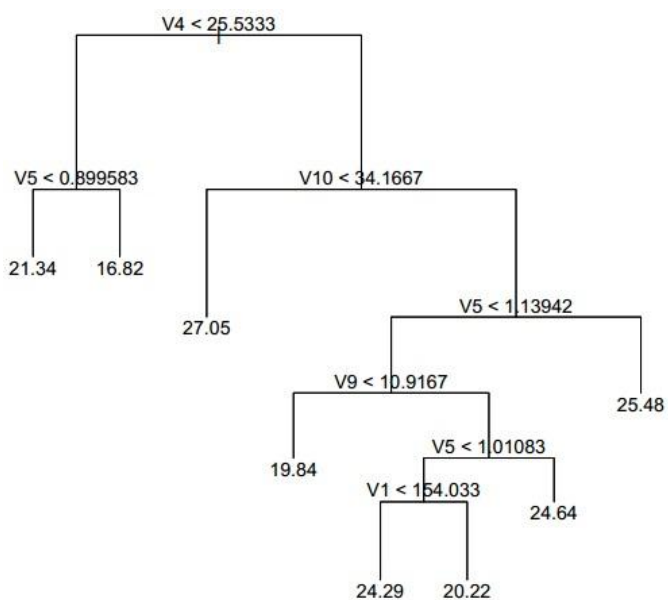


Figura 25. Árvore de decisão para capacidade de expansão e variáveis morfológicas em milho-pipoca, safra de 2021, sob condição hídrica WW.

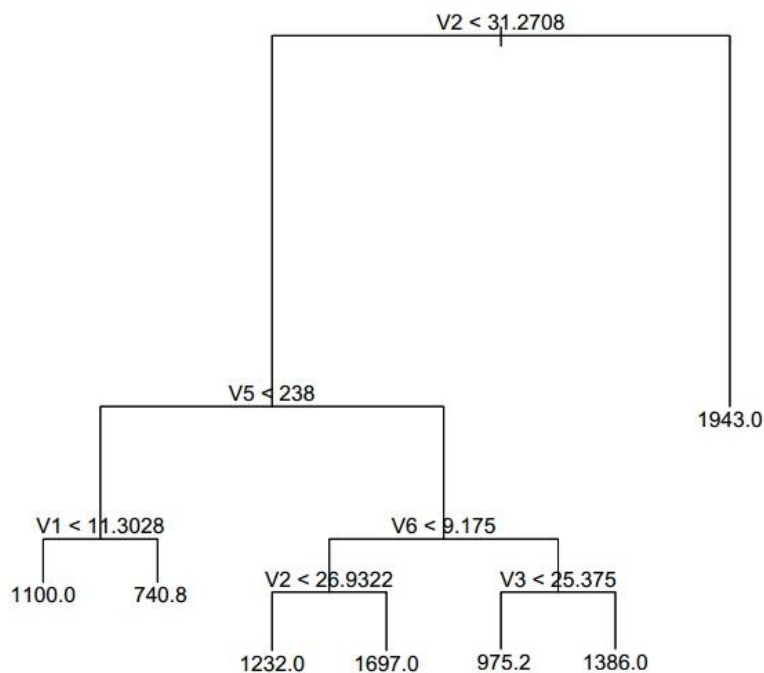


Figura 26. Árvore de decisão para GY e variáveis agrônômicas em milho-pipoca, safra de 2021, sob condição hídrica WW.

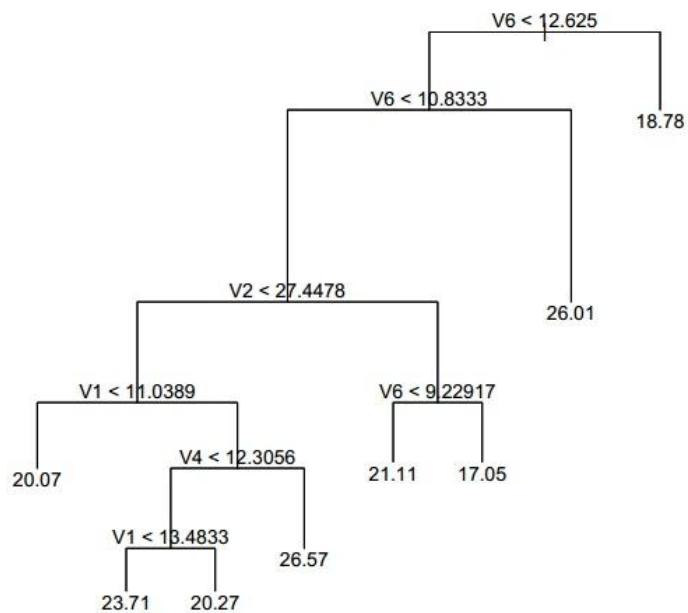


Figura 27. Árvore de decisão para PE e variáveis agrônômicas em milho-pipoca, safra de 2021, sob condição hídrica WW.

Safra 2021, na condição WS.

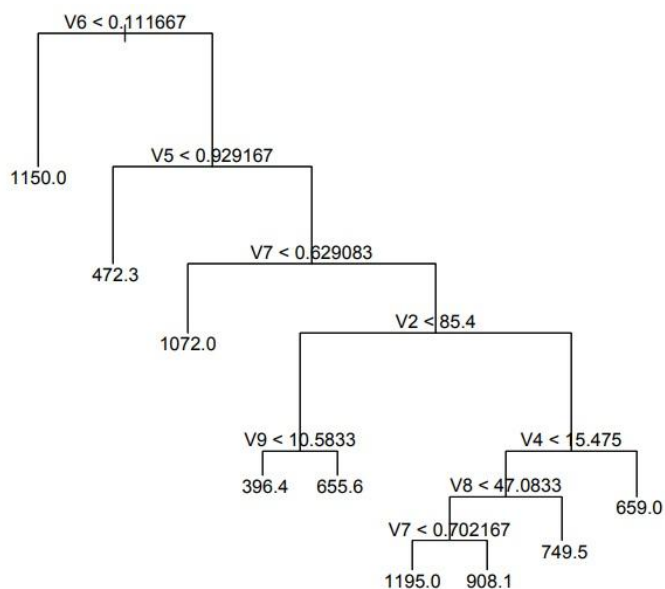


Figura 28. Árvore de decisão para GY e variáveis morfológicas em milho-pipoca, safra de 2021, sob condição hídrica WS.

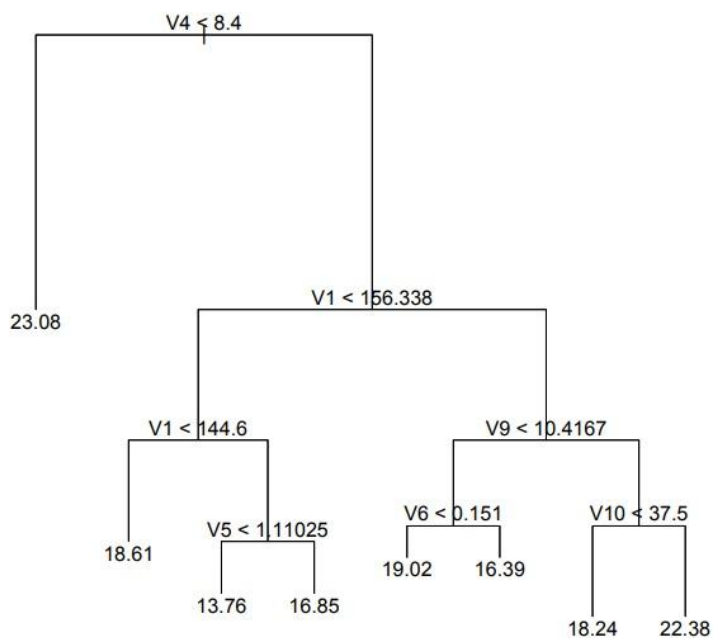


Figura 29. Árvore de decisão para capacidade de expansão e variáveis morfológicas em milho-pipoca, safra de 2021, sob condição hídrica WS.

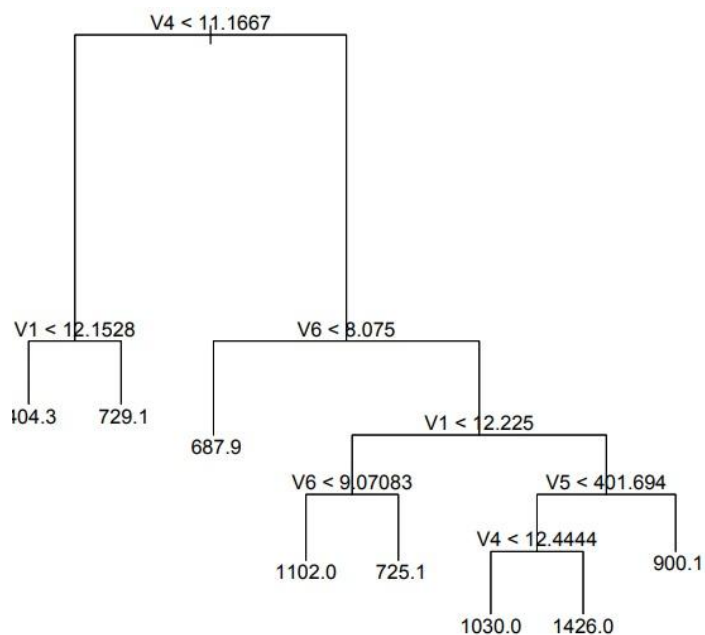


Figura 30. Árvore de decisão para GY e variáveis agronômicas em milho-pipoca, safra de 2021, sob condição hídrica WS.

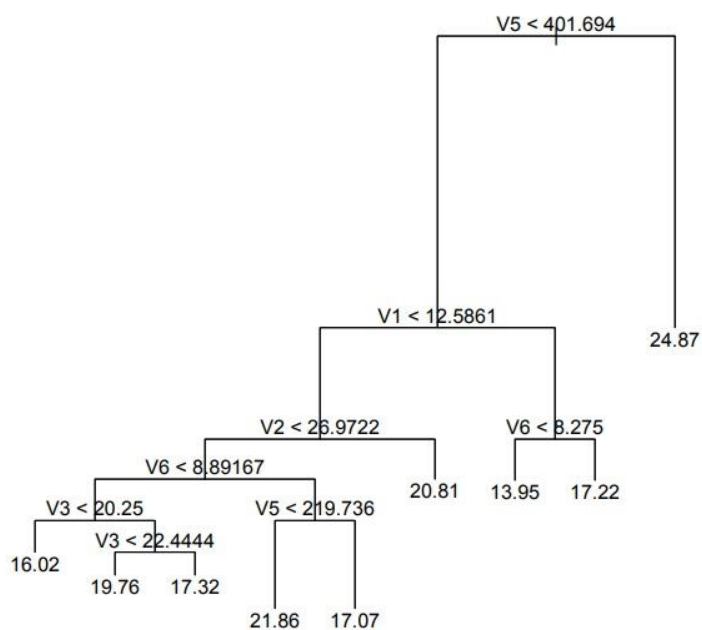


Figura 31. Árvore de decisão para PE e variáveis agronômicas em milho-pipoca, safra de 2021, sob condição hídrica WS.