



**VITÓRIO ANTÔNIO PEREIRA DE SOUZA**

**SELEÇÃO INDIRETA PARA O NÚMERO DE GRÃOS POR  
VAGEM E TAMANHO DOS GRÃOS VISANDO A  
PRODUTIVIDADE EM SOJA**

**LAVRAS-MG**

**2023**

**VITÓRIO ANTÔNIO PEREIRA DE SOUZA**

**SELEÇÃO INDIRETA PARA O NÚMERO DE GRÃOS POR VAGEM E TAMANHO  
DOS GRÃOS VISANDO A PRODUTIVIDADE EM SOJA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

Prof. Dr. Adriano Teodoro Bruzi  
Orientador

**LAVRAS-MG**

**2023**

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

Souza, Vítório Antônio Pereira de.

Seleção indireta para o número de grãos por vagem e tamanho dos grãos visando a produtividade em soja / Vítório Antônio Pereira de Souza. - 2023.

53 p.

Orientador(a): Adriano Teodoro Bruzi.

Dissertação (mestrado acadêmico) - Universidade Federal de Lavras, 2023.

Bibliografia.

1. Glycine max (L) Merrill. 2. Melhoramento Genético. 3. Associação entre caracteres. I. Bruzi, Adriano Teodoro. II. Título.

**VITÓRIO ANTÔNIO PEREIRA DE SOUZA**

**SELEÇÃO INDIRETA PARA O NÚMERO DE GRÃOS POR  
VAGEM E TAMANHO DOS GRÃOS VISANDO A  
PRODUTIVIDADE EM SOJA**

**INDIRECT SELECTION FOR THE NUMBER OF BEANS PER POD AND BEAN  
SIZE AIMING AT SOYBEAN PRODUCTIVITY**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

APROVADO em 31 de julho de 2023.

Dra. Christiane Augusta Diniz Melo  
Dr. Anderson Dallastra

UFLA  
GDM

Prof. Dr. Adriano Teodoro Bruzi  
Orientador

**LAVRAS-MG  
2023**

*A Deus e a minha família pelo apoio incondicional em todas as etapas da minha vida.*

*Dedico*

## AGRADECIMENTOS

À Deus, pela vida, saúde, força e por, mesmo nos momentos de incertezas, ter me mantido focado.

Aos meus pais, Antônio Bispo e Auda Suely e à minha irmã, Maria Vitória pela força, pelos conselhos, pela confiança e por todo o amor que me proporcionaram trilhar esse caminho.

À minha noiva, Arianna Menezes pela confiança, amor, paciência, conselhos e por ser minha companheira em todos os momentos.

Ao meu orientador Adriano Teodoro Bruzi pelos ensinamentos acadêmicos e de vida e por toda a confiança transmitida.

À equipe Pesquisa Soja, em especial Carlos Henrique, Mateus Piza, Taine Rocha e demais colegas e amigos que contribuíram diretamente ou indiretamente com o desenvolvimento do trabalho e minha formação acadêmica.

À Universidade Federal de Lavras (UFLA), especialmente ao Departamento de Biologia (DBI) e ao Programa de Pós-graduação em Genética e Melhoramento de Plantas pela oportunidade.

Aos membros da banca pela disponibilidade e contribuições ao trabalho.

À empresa GDM pela parceria e apoio no desenvolvimento desse trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), pela concessão da bolsa de estudos.

Às demais agências de fomento, pelo financiamento concedido à realização deste projeto.

Por fim, agradeço a todos que se fizeram presente no desenvolvimento desse trabalho e em minha jornada acadêmica.

Muito obrigado!

## RESUMO

Em soja a produtividade potencial de grãos é dependente do número de plantas por área, número de vagens por planta, número de grãos por vagem e do peso médio dos grãos. Devido a importância dessa cultura no cenário mundial, é necessário constantemente investir em novas tecnologias de melhoramento genético e de manejo, visando aumentar os ganhos em produtividade de grãos. Assim, objetivou-se estudar os parâmetros genéticos, fenotípicos e a correlação entre os caracteres peso de mil grãos, número de grãos por vagem, área foliar e produtividade de grãos em soja e identificar qual estratégia de seleção é mais eficiente para incrementar a produtividade de grãos. Para obtenção da população segregante foram utilizadas duas cultivares contrastantes para os caracteres tamanho de grãos e número de grãos por vagem. A população F<sub>2</sub> foi conduzida em Bulk. Na maturação fisiológica, foram colhidas plantas individuais, que foram classificadas de acordo com o rendimento de peneiras. As progênies F<sub>2:3</sub> e F<sub>2:4</sub> foram conduzidas e avaliadas nos anos agrícolas 2021/22 e 2022/23. Avaliaram-se os caracteres: dias para o florescimento, tipo de folíolo, maturação absoluta, altura de planta, altura de inserção da primeira vagem, índice de acamamento, número de vagens por planta, número de grãos por vagem, peso de mil grãos e produtividade de grãos. As análises estatísticas foram realizadas com o auxílio do ambiente estatístico R, utilizando uma abordagem de modelos mistos (REML/BLUP). A contribuição dos componentes de produção para a produtividade de grãos variou em função da progênie, evidenciando que nem sempre o genótipo de maior equivalência entre os componentes é o de maior produtividade. O peso de mil grãos e o número de grãos por vagem apresentaram relação inversa. Essa relação indica que, quando os grãos são maiores, isto é, possuem maior peso individual, o número de grãos por vagem tende a ser menor, e vice-versa. Assim, um aumento no tamanho dos grãos pode levar a uma diminuição no número total de grãos formados por planta. A seleção indireta, considerando os caracteres número de grãos por vagem ou tamanho de grãos, não foi a melhor estratégia para aumentar a produtividade na cultura da soja, devido à complexa interação entre os diferentes componentes de produção. Assim, a melhor abordagem para selecionar linhagens superiores de soja é focar diretamente na produtividade de grãos, pois esta estratégia contempla a interação que existe entre os diferentes fatores que afetam a produtividade e, por consonância, permite a identificação de genótipos superiores mais eficazmente.

**Palavras-chave:** *Glycine max* (L.) Merrill. Melhoramento Genético. Rendimento de peneira. Associação entre caracteres. Componentes de produção.

## ABSTRACT

In soybeans, potential grain productivity is dependent on the number of plants per area, number of pods per plant, number of grains per pod and the average weight of the grains. Due to the importance of this crop on the world stage, it is necessary to constantly invest in new genetic improvement and management technologies, aiming to increase gains in grain productivity. Thus, the objective was to study the genetic and phenotypic parameters and the correlation between the characters weight of a thousand grains, number of grains per pod, leaf area and grain productivity in soybeans and identify which selection strategy is most efficient to increase soybean productivity grains. To obtain the segregating population, two contrasting cultivars were used for the characters grain size and number of grains per pod. The F<sub>2</sub> population was conducted in Bulk. At physiological maturity, individual plants were harvested and classified according to sieve yield. The F<sub>2:3</sub> and F<sub>2:4</sub> progenies were conducted and evaluated in the 2021/22 and 2022/23 agricultural years. The following characters were evaluated: days to flowering, type of leaflet, absolute maturation, plant height, insertion height of the first pod, lodging index, number of pods per plant, number of grains per pod, weight of a thousand grains and grain productivity. Statistical analyzes were performed with the aid of the R statistical environment, using a mixed model approach (REML/BLUP). The contribution of the production components to grain productivity varied depending on the progeny, showing that the genotype with the greatest equivalence between the components is not always the one with the highest productivity. The weight of a thousand grains and the number of grains per pod showed an inverse relationship. This relationship indicates that, when the grains are larger, that is, they have greater individual weight, the number of grains per pod tends to be smaller, and vice versa. Thus, an increase in grain size can lead to a decrease in the total number of grains formed per plant. Indirect selection, considering the characters number of grains per pod or grain size, was not the best strategy to increase productivity in soybean cultivation, due to the complex interaction between the different production components. Therefore, the best approach to selecting superior soybean lines is to focus directly on grain productivity, as this strategy takes into account the interaction that exists between the different factors that affect productivity and, accordingly, allows the identification of superior genotypes more effectively.

**Keywords:** *Glycine max* (L.) Merrill. Plant breeding. Sieve yield. Association between Traits. Grain Yield Components.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>10</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> .....	<b>12</b>
<b>2.1</b>	<b>A cultura da soja no brasil</b> .....	<b>12</b>
<b>2.2</b>	<b>Componentes de produção em soja</b> .....	<b>13</b>
<b>2.3</b>	<b>Associação entre caracteres na cultura da soja</b> .....	<b>15</b>
<b>2.4</b>	<b>Emprego da análise de imagens no melhoramento de plantas</b> .....	<b>16</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b> .....	<b>19</b>
<b>3.1</b>	<b>Obtenção das progênies</b> .....	<b>19</b>
<b>3.2</b>	<b>Condução dos experimentos</b> .....	<b>19</b>
<b>3.4</b>	<b>Caracteres avaliados</b> .....	<b>21</b>
<b>3.5</b>	<b>Uso de imagens na fenotipagem</b> .....	<b>22</b>
<b>3.6.1</b>	<b>Análise dos dados</b> .....	<b>23</b>
<b>3.6.2</b>	<b>Resposta a seleção</b> .....	<b>26</b>
<b>3.6.3</b>	<b>Análise geométrica</b> .....	<b>27</b>
<b>3.6.4</b>	<b>Análise de agrupamento</b> .....	<b>28</b>
<b>3.6.5</b>	<b>Aspectos computacionais</b> .....	<b>30</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b> .....	<b>31</b>
<b>5</b>	<b>CONCLUSÕES</b> .....	<b>43</b>
	<b>REFERÊNCIAS</b> .....	<b>44</b>
	<b>ANEXOS</b> .....	<b>50</b>

## 1 INTRODUÇÃO

Em soja, a produtividade potencial de grãos é dependente do número de plantas por área, número de vagens por planta, número de grãos por vagem e do peso médio dos grãos (VOGEL *et al.*, 2021). Com o objetivo de aumentar a produção, a principal estratégia é incrementar a produtividade, devido à sustentabilidade e a dificuldade em se abrir novas áreas. Diante disto, o caráter produtividade é o de maior interesse nos programas de melhoramento genético vegetal (BORÉM, 2021).

Um questionamento que se concerne então, refere-se aos componentes de produção, isto é, qual seria a relação ideal entre o número de vagens por planta (X), número de grãos por vagem (Y) e peso de mil grãos (Z) visando a maximização da produtividade. Considerando que a produtividade (W) é o produto de X, Y e Z, ou seja, três eixos, Grafius (1964) fez uma analogia da produtividade como sendo o volume de um paralelepípedo. O autor realçou que o paralelepípedo com maior volume é o cubo, ou seja, a figura em que os três eixos contribuem igualmente para o volume.

Os caracteres número de grãos por vagem (NGV), número de vagens por planta (NVP), rendimento de peneira (RP) e peso de mil grãos (PMG) são características que podem ser indicativos da produtividade de grãos. Como a produtividade de grãos é uma característica controlada por muitos genes e possui baixa herdabilidade os programas de melhoramento têm buscado caracteres com controle genético mais simples, e que possuem alta correlação com a produtividade visando à seleção indireta. Contudo, relatos na literatura evidenciam, também, que as correlações entre X (NVP) e Y (NGV), X (NVP) e Z (PMS) ou Y (NGV) e Z (PMS) são normalmente negativas, ou seja, o crescimento de um deles repercute na redução dos demais (VOGEL *et al.*, 2021; COELHO *et al.*, 2002; LANA *et al.*, 2003).

Por outro lado, Dallastra *et al.* (2014) comentaram que caracteres como número de ramos por planta, número de vagens por planta e número de sementes por planta fazem parte dos componentes de produção, correlacionando-se positivamente à produtividade, pois quanto maior o valor destes, maior será a magnitude da produtividade de grãos de genótipos de soja.

No que tange a seleção indireta, a área dos trifólios é outra característica que pode estar relacionada com a produtividade de grãos. Sayama *et al.* (2017), estudando a relação entre o número de grãos por vagem e o tipo de folíolo em soja, constataram que o gene *Ln* da soja controla de forma pleiotrópica as características número de grãos por vagem e tipo de folíolo. Sendo o alelo dominante *Ln* responsável por folíolos largos e alelo o recessivo *ln* responsável por folíolos estreitos. Como existe efeito pleiotrópico para as características NGV

e tipo de folíolo espera-se que cultivares com folíolo lanceolado tenham em média um maior número de grãos por vagem.

A determinação da área foliar tem sido obtida pela mensuração direta da largura e comprimento da folha obtidos manualmente. Este é um método de baixa precisão devido à discrepância da qualificação entre os avaliadores, na calibração dos equipamentos utilizados no processo, e principalmente, devido ao cansaço físico e mental dos avaliadores (SILVA *et al.*, 2019). Muito embora, a fenotipagem manual ainda seja a mais utilizada na maioria dos programas de melhoramento genético de soja no Brasil, esta técnica possui menor precisão que as ferramentas de fenotipagem de alto rendimento. A análise de imagens digitais tem sido usada para mensuração de caracteres morfológicos e fisiológicos em centenas ou milhares de plantas em um curto intervalo de tempo e com baixa exigência em mão de obra (JARQUIN, 2019; MAIMAITIJIANG *et al.*, 2019). As imagens digitais têm sido empregadas para a determinação da cobertura de dossel vegetativo em soja ou da radiação solar interceptada pela cultura (GASPAR; CONLEY, 2015; XAVIER *et al.*, 2017).

Neste sentido, Li *et al.* (2021) relataram que a expansão das tecnologias está propiciando um importante avanço na fenotipagem de alto rendimento com o aumento na aquisição e processamento de dados de várias culturas, o que pode reduzir muito o trabalho manual e o tempo necessário para obter informações fenotípicas da cultura. A fenotipagem de alto rendimento, além de ser mais precisa, possibilita a avaliação de mais genótipos de maneira mais rápida (ARAUS *et al.*, 2022).

Diante deste cenário, objetivou-se estudar os parâmetros genéticos, fenotípicos e a correlação entre os caracteres tamanho de grãos, número de grãos por vagem, área foliar (obtida pelo processamento de imagem) e a produtividade de grãos em soja, bem como identificar entre as estratégias de seleção direta e indireta, qual é a mais eficiente para o incremento de produtividade de grãos nos programas de melhoramento da soja.

## 2 REFERENCIAL TEÓRICO

### 2.1 A cultura da soja no Brasil

A soja (*Glycine max* (L.) Merrill) é uma das culturas mais importantes do mundo e é a principal cultura do Brasil, que é o maior produtor e exportador (CONAB, 2023). A cultura é originária do nordeste da china, em uma região conhecida como Manchúria, onde consta indícios de sua domesticação, sendo posteriormente disseminada para os Estados Unidos, Europa e América do Sul (ARANTES; SOUZA, 1993).

No Brasil, a soja foi introduzida na Bahia, em 1882, nesse período a cultura não se desenvolveu na região devido ao fato dos materiais introduzidos não serem adaptados às condições de baixa latitude do estado, em 1908, imigrantes japoneses introduziram a soja em São Paulo e no Rio Grande do Sul, onde a cultura teve sucesso, apresentando uma melhor adaptação, uma vez que as condições climáticas são mais semelhantes às das regiões tradicionais de cultivo (SILVA *et al.*, 2022). A partir da década de 1970 as lavouras de soja ganharam o território nacional, principalmente no Cerrado brasileiro onde se concentrou as principais fronteiras agrícolas, e hoje a cultura é produzida de norte a sul do país. O melhoramento genético teve papel fundamental nessa expansão, devido a descoberta de genótipos com período juvenil longo, que demonstraram boa adaptação a condições de baixa latitude brasileira (FARIAS; NEPOMUCENO; NEUMAIER, 2007; MARTINELLI *et al.*, 2017).

O Brasil é o maior produtor de soja do mundo, o país produziu na safra 2022/23, 155,736 milhões de toneladas em uma área plantada de 44,031 milhões de hectares, com produtividade média de 3500 kg ha<sup>-1</sup>. Os Estados Unidos ficam com a segunda colocação com uma produção de 119,88 milhões de toneladas em uma área plantada de 35,450 milhões de hectares e produtividade média de 3384 kg ha<sup>-1</sup> (USDA, 2022). A Argentina é o terceiro maior produtor e destaca-se no cenário mundial.

No Brasil na safra 22/23 a cultura teve um aumento de 6,1% na área plantada comparado ao da safra 2021/22. A produção obtida foi 24% maior do que na safra 2021/22, e a produtividade média alcançada foi de 3537 kg ha<sup>-1</sup>, com produtividades recordes em vários estados, sendo essa produtividade 16,9% maior que a alcançada na safra anterior. Do total produzido, 75% estão nas regiões Centro-Oeste e Sul. O estado do Mato Grosso lidera o ranking como maior produtor brasileiro de soja com uma produção de 45,600 milhões de toneladas em uma área plantada de 12,086 milhões de hectares, representando 27,4% da área

plantada, seguido por Paraná com 22,384 milhões de toneladas produzidas, Goiás com 17,734 milhões de toneladas e Rio grande do Sul com 14,513 milhões de toneladas. A produção de soja tem crescido bastante na região nordeste do Brasil, que na última safra produziu 15,131 milhões de toneladas, com destaque para o estado da Bahia que teve a maior produtividade do país com 4020 kg ha<sup>-1</sup> (CONAB, 2023).

O segmento sojícola tem grande contribuição no desenvolvimento socioeconômico do país, a cadeia produtiva da soja teve grande participação na balança comercial do Brasil, as estimativas para exportações do grão, farelo e óleo, sugerem uma receita de 66,8 bilhões de dólares, para a safra 2022/2023 tornando a soja o principal item de exportação do país (ABIOVE, 2023), além disso, estima-se que para cada emprego direto com soja são gerados 12,66 empregos indiretos (MTE, 2010).

O custo relativamente baixo de produção, em comparação a outras culturas, a alta qualidade da proteína e a produção de óleo vegetal de qualidade fazem da soja importante matéria-prima para produção de óleo, biodiesel, lubrificantes, farelo proteico para alimentação humana e animal, adubo verde etc. (SILVA *et al.*, 2022). Em princípio, os teores de óleo e proteínas das sementes de soja são determinados geneticamente, porém, fortemente influenciados pelo ambiente, principalmente durante o período de enchimento dos grãos.

## 2.2 Componentes de produção em soja

Em soja, assim como em outras culturas, o rendimento de grãos é uma característica complexa, que depende de muitas variáveis (MALIK *et al.*, 2007). O número de vagens por planta e grãos por vagem são os principais determinantes da produtividade de grãos em soja (NAVARRO JUNIOR; COSTA, 2002; KANTOLIC *et al.*, 2013).

O desenvolvimento reprodutivo em culturas de grãos é iniciado pela floração, onde o número de frutos e sementes que serão produzidos é determinado durante este período. A produção de flores e sementes é considerada uma fase chave para definir os componentes de produção de uma cultura (EGLI, 2005).

A produtividade da soja é construída durante todo o ciclo da cultura e envolve vários fatores como: histórico da área, preparo do solo, plantio, qualidade das sementes, adubação, disponibilidade hídrica, manejo de plantas daninhas, pragas e doenças e a qualidade da colheita. O manejo adequado desses componentes é fundamental para garantir uma boa produtividade de soja de alta qualidade. Todos esses fatores vão afetar a produtividade da soja. Então para ter uma noção do esperado para produtividade antes mesmo da colheita

existem algumas características que são chamadas de componentes de produção que afetam diretamente a produtividade das plantas, são eles: número de plantas por área, número de vagens por planta, número de grãos por vagem e peso dos grãos, que normalmente é medido através do peso de mil grãos (PMG) (RICHETTI, 2021).

O número de vagens por planta é um fator importante na determinação da produtividade potencial da cultura. Esse número pode variar dependendo de uma variedade de fatores, como genética, ambiente, associado a ocorrência de fatores de cunho biótico e abiótico, atrelado a prática de manejo. Geralmente, as plantas de soja produzem entre 50 a 100 vagens por planta, mas isso pode variar de acordo com a cultivar de soja e as condições de cultivo. Fatores como fertilidade do solo, temperatura, umidade e pressão de pragas e doenças podem afetar o número de vagens que uma planta produz (SIDYAMA *et al.*,2015).

As plantas de soja geralmente produzem vagens em ráculos, com cada ráculo contendo tipicamente de 3 a 5 vagens. Em geral, plantas de soja com mais ramos produzirão mais vagens por planta, e plantas mais espaçadas terão mais ramos e, portanto, mais vagens. Vale a pena notar que, embora o número de vagens por planta seja um fator importante na determinação do potencial da produtividade, não é o único fator. Outros fatores, como tamanho, peso e qualidade geral do grão, também podem afetar a produtividade (SIDYAMA *et al.*,2015).

Sayama *et al.* (2017) mostraram que o gene *Ln* da soja controla de forma pleiotrópica as características número de grãos por vagem e tipo de folíolo, sendo o gene recessivo *Ln* responsável por conferir o fenótipo folíolos estreitos, com número de grãos por vagem significativamente maior do que a linhagem com folíolos largos, que fenotipicamente apresenta um menor número de grãos por vagens.

O tamanho da semente de soja é uma característica importante que pode afetar o potencial da produtividade e a qualidade fisiológica. As sementes de soja geralmente variam entre uma faixa de tamanho típica de 5 a 8 milímetros. O tamanho da semente pode variar dependendo de vários fatores, incluindo genética, condições ambientais e práticas de manejo. Além do tamanho, outros fatores como cor, forma e textura da semente também são importantes na determinação da qualidade e comercialização. No geral, o tamanho da semente de soja é um fator importante a ser considerado na produção de soja, mas é apenas um dos muitos fatores que podem afetar o potencial de rendimento e a qualidade da semente.

Grafius (1964) propôs avaliar a contribuição dos componentes de produção na produtividade de grãos e, também, na capacidade do genótipo em modificar seu

desenvolvimento, morfologia ou comportamento com base nas condições em que se encontra (plasticidade fenotípica) a partir de premissas geométricas. Segundo ele, a produtividade de grãos pode ser representada pelo volume de um paralelepípedo e os eixos da referida figura geométrica são os três componentes primários, número de grãos por vagem, número total de vagem por planta e peso de mil grãos. O autor ainda afirma que a figura geométrica de maior volume representaria o ideótipo para produtividade, essa situação é encontrada quando os três componentes primários apresentarem a mesma contribuição. A produtividade de grãos é um caráter complexo e, portanto, de difícil seleção; o seu aumento depende da integração de várias técnicas de melhoramento e pós-melhoramento

### **2.3 Associação entre caracteres na cultura da soja**

O sucesso da soja no Brasil deve-se a diversos fatores, com destaque ao melhoramento genético da espécie, com o desenvolvimento de cultivares mais produtivas e adaptadas as variações edafoclimáticas e resistente a diversas doenças. A produtividade de grãos é a principal característica a ser melhorada nas empresas, contudo, embora esse seja o caráter mais estudado ainda é difícil entender a sua herança e as implicações em outras características, isso porque a produtividade de grãos é um caráter de natureza quantitativa, ou seja, é controlado por muitos genes. Por isso é muito influenciado pelo ambiente (baixa herdabilidade) e apresenta distribuição contínua, ou seja, as frequências da estimativa assumem uma curva normal (SILVA *et al.*, 2017).

O fenótipo é definido como as diferentes manifestações de um dado caráter. Este é influenciado pela natureza genética de seus parentais e pelo ambiente que pode ser definido por todas as condições que afetam o crescimento e desenvolvimento do organismo. Dessa forma, o fenótipo (F) a ser obtido será o resultado da ação do genótipo (G), do ambiente (A), e da interação de ambos (GA), ou seja,  $F=G+A+GA$  (RAMALHO *et al.*, 2012). Meotti *et al.*, (2012) observaram o efeito da interação GA sobre os componentes de produção da soja em um estudo objetivando avaliar o efeito de épocas de semeadura no desempenho agrônômico de cultivares de soja.

Os programas de melhoramento visam não somente aprimorar um caráter principal, mas também melhorar o maior número possível de caracteres. Assim, conhecer as informações a respeito da correlação entre as características que se deseja trabalhar é de grande interesse para os melhoristas (MIRANDA, 2006). Dessa forma, o melhorista quase sempre deve trabalhar com vários caracteres ao mesmo tempo. Fazer isso com eficiência é o

grande desafio desses profissionais, isso porque a maioria dos caracteres estão relacionados, muitas vezes em direções opostas (RAMALHO *et al.*, 2012).

Segundo Ramalho *et al.* (2012), correlação é a medida de intensidade de associação linear de duas variáveis; pode ser positiva ou negativa. Coeficientes de correlação são adimensionais, oscilando positiva ou negativamente de 0 a 1, em que correlação nula evidencia não ausência e falta de relação linear entre características (NOGUEIRA, 2012) e coeficientes de correlação situados de -0,5 a -1 e de 0,5 a 1 são valorizados devido a magnitude dos valores na interpretação aplicada das correlações (LOPES *et al.*, 2002).

Por meio da correlação de caracteres é possível verificar se a variação de um determinado caráter está associada à de outro. A correlação pode ser dividida em fenotípica, genotípica e ambiental, sendo os dois últimos os principais, pois na determinação do fenótipo estão envolvidas tanto as causas genéticas quanto as ambientais. Considerando dois caracteres correlacionados fenotipicamente e de baixa herdabilidade, a correlação de ambiente predomina na correlação fenotípica, e se for de alta herdabilidade a genotípica se destaca (MIRANDA, 2006; FALCONER, 1987; CRUZ *et al.*, 2012).

O conhecimento da associação de caracteres de fácil mensuração e de alta herdabilidade com a produtividade é de interesse e de extrema importância quando o objetivo é realizar a seleção indireta. Isso porque se existe correlação a seleção pode ser realizada para a característica de alta herdabilidade e fácil mensuração, aumentando a eficiência da seleção e consequentemente o sucesso do programa de melhoramento e desenvolvimento da cultura (LEITE *et al.*, 2015).

A interação alélica refere-se ao fenômeno em que diferentes alelos, ou versões de um gene, interagem entre si para determinar a expressão de uma característica específica. O número de grãos por vagem e tamanho do grão de soja é uma característica complexa que é influenciada por múltiplos genes, cada um com diferentes alelos de características quantitativas (QTLs) que interagem entre si de maneiras complexas. A utilização de metodologias que possibilitam melhorias na seleção de genótipos superiores é importante para aumentar a eficiência dos programas de melhoramento, como é o caso da análise de imagens, que tem sido amplamente utilizada na seleção das mais diversas características (CARNEIRO, 2018).

## **2.4 Emprego da análise de imagens no melhoramento de plantas**

Entre um dos fatores que contribuíram para que as previsões de Malthus não tenham

ainda se concretizado é o melhoramento genético vegetal. No entanto, embora atualmente em vários países a produção de alimentos é exorbitante, em diversas regiões do mundo ainda se encontram subjugadas à fome e a desnutrição (BORÉM, 2021), figurando-se necessário a progressividade na eficiência dos programas de melhoramento. A fenotipagem de alto rendimento surge como uma ferramenta, visando tais ganhos. (ARAUS *et al.*, 2018). Apesar de ser capaz de caracterizar as respostas da cultura ao ambiente para muitas características a fenotipagem tradicional de plantas está sujeita a inúmeros problemas, tanto na coleta como no processamento desses dados.

Neste contexto a fenotipagem de alto rendimento pode auxiliar na melhor coleta dos dados melhorando a precisão, acurácia, velocidade e custo dos experimentos. A fenotipagem de alto rendimento é uma ferramenta que os melhoristas podem utilizar para avaliar mais indivíduos de maneira mais precisa, proporcionando um aumento na eficiência dos programas de melhoramento (ARAUS *et al.*, 2018). Nos últimos anos, com o avanço das ciências correlatas à fenotipagem de alto rendimento permitiu o surgimento de novos métodos de fenotipagem de plantas, como por exemplo a obtenção de imagens por câmeras hiper espectrais. O emprego da fenotipagem nos experimentos proporciona avaliações mais rápidas e precisas quando comparadas aos métodos convencionais (SOUSA *et al.*, 2015).

Devido a sua importância socioeconômica, a soja é uma cultura alvo de diversos estudos, principalmente com novas tecnologias, como o desenvolvimento de novos eventos de transgenia e estudos com edição gênica. Roth *et al.*, (2022) afirmaram que a cultura da soja pode ser considerada a cultura modelo para testar novos métodos de fenotipagem.

Características diretamente correlacionadas com a produtividade de grãos da cultura da soja, como número de vagens por planta e número de grãos por planta são mensuradas manualmente, dessa maneira, deve-se levar em consideração a baixa precisão associada à fenotipagem manual devido à discrepância de qualificação entre os avaliadores, calibração dos equipamentos utilizados no processo e principalmente devido ao cansaço físico e mental dos avaliadores, além do número limitado de genótipos que podem ser avaliados, tornando-o extremamente ocioso. Neste sentido, a fenotipagem automatizada pode substituir a fenotipagem manual dos componentes de produção da soja, melhorar a precisão e a velocidade do processamento de dados de campo e ajudar o melhorista a identificar características de interesse para a seleção de genótipos superiores.

As imagens consistem em informações visuais armazenadas em forma digital. Estas são representadas por matrizes numéricas bidimensionais, onde cada ponto é um pixel que

assume valores entre 0 (preto) e 255 (branco). As imagens são classificadas em binárias, escala de cinza e coloridas. Para cada um destes valores existe um tom de cinza associado. Em ambas, a imagem é constituída por uma única matriz de pixels. Diferente das imagens binárias, as imagens coloridas são mais complexas, diferindo na quantidade de matrizes que as compõe, que depende do sistema de cor adotado para representá-las (RODRIGUES, 2022).

Existem diferentes sistemas de cores, dos quais, o mais usual é o RGB, composto por três canais: R – vermelho; G – Verde; B – Azul. Cada pixel é composto pela combinação da intensidade destes três canais, que proporciona a constituição de uma única cor que observamos visualmente (KALAIVANI *et al.*, 2013). Outros sistemas de cores são utilizados para representar imagens coloridas. Dentre estes, o sistema HSV (hue, saturation, value) e os propostos pela Comissão Internacional de Iluminação (Comission Internationale de l'Eclairage - CIE) têm sido muito utilizados para auxiliar no processamento de imagens digitais. É ampla a aplicação de imagens RGB, em trabalhos visando a identificação de plantas daninhas (LOPES *et al.*, 2023), manejo de doenças (TREVISAN, 2022) e em programas de melhoramento em culturas como a soja (FERREIRA, 2022).

As análises de imagens podem ser subdivididas nas seguintes etapas: aquisição, pré-processamento, segmentação e extração de características. A aquisição de imagens é uma etapa crucial na utilização de imagens digitais em pesquisas científicas, uma vez que será este momento que irá definir a qualidade das imagens a serem utilizadas. Existem câmeras que possuem capacidades diferenciadas na captação de informações, como térmicas ou multiespectrais e outras relativamente mais simples que possuem preço mais acessível, por isso são comumente utilizadas (TATTARIS *et al.*, 2016). A nível de laboratório, algumas pesquisas têm sido realizadas com câmeras acopladas em estúdios fotográficos. Já na fenotipagem em campo são empregadas câmeras acopladas a Aeronave Remotamente Pilotada (ARP) ou plataformas específicas para fenotipagem (WALTER *et al.*, 2015).

### 3 MATERIAL E MÉTODOS

#### 3.1 Obtenção das progênies

Para obtenção da população segregante, foram utilizados dois genitores: IAC-100 e BRASMAX BÔNUS IPRO - 8579 RSF. Os genitores foram escolhidos considerando o número de grãos por vagem e o tamanho dos grãos. O genitor IAC-100 é uma cultivar convencional desenvolvida pelo Instituto Agrônomo de Campinas (IAC). Pertence ao grupo de maturidade relativa 8.1, apresenta hábito de crescimento determinado, folhas lanceoladas, moderada resistência a percevejos, alta prolificidade (podendo apresentar em média até 4,5 grãos por vagem) além do bom desempenho produtivo. A cultivar BRASMAX BÔNUS IPRO - 8579 RSF, desenvolvida pela empresa GDM e comercializada pela empresa BRASMAX, pertence ao grupo de maturidade relativa 7.9, recomendada para região central do Brasil, sudeste, norte e nordeste. A cultivar apresenta hábito de crescimento indeterminado e folhas ovaladas. Além disso, apresenta classificação de peneira 7 mm e resistência ao cancro da haste.

Os cruzamentos foram realizados em casa de vegetação na estação de Cria da GDM Seeds localizada em Porto Nacional - TO durante a safra 2019/2020. A hibridação manual seguiu procedimento semelhante ao descrito por Borém *et al.* (2009). As sementes da geração  $F_1$  foram multiplicadas na safra de inverno de 2020 em casa de vegetação, em Porto Nacional, adotando duas plantas por vaso. A população segregante  $F_2$  foi conduzida utilizando o método Bulk. Esta foi multiplicada na safra verão 2020/2021.

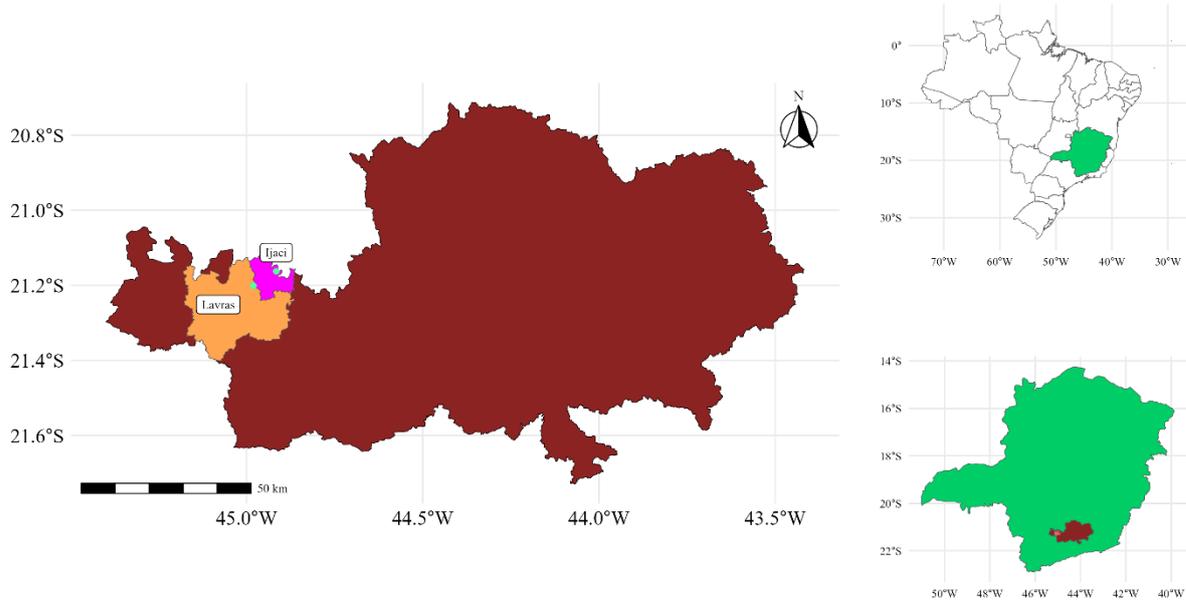
Na maturação fisiológica foram colhidas plantas individuais visando à obtenção das progênies  $F_{2:3}$ . As plantas foram colhidas e trilhadas individualmente utilizando debulha manual, os grãos foram quantificados e acondicionados em sacos de papel. Posteriormente, foi realizada a classificação desses grãos quanto ao tamanho de peneira, classificando as progênies em três grupos, peneira 5.5 mm, 6.0 mm e 7.0 mm.

#### 3.2 Condução dos experimentos

Os experimentos foram conduzidos em dois locais, nos municípios de Lavras-MG, no Centro de Desenvolvimento Científico Tecnológico em Agropecuária - Fazenda Muquém, situada à latitude de 21°14' S, longitude 45°00' W e altitude de 918 m, e no município de Ijaci-MG, Centro de Desenvolvimento Científico e Tecnológico em Agropecuária – Fazenda

Palmital, situada à latitude de 21°09' S, longitude 44°54' W e altitude de 920 m (Figura 1), durante os anos agrícolas 2021/2022 e 2022/2023.

Figura 1 – Representação dos locais de implementação dos experimentos no Campo das Vertentes, Minas Gerais – Brasil.



Fonte: Do autor (2023).

Na safra 2021/22, foram avaliadas as progênies  $F_{2:3}$  em parcelas constituídas de uma linha de um metro espaçadas em 0,6 metro, e a densidade de semeadura foi ajustada conforme o grupo de maturidade absoluta das progênies que foi em média de 8,0, estimado na safra 2020/21, na geração  $F_2$ . Os experimentos com progênies de diferentes tamanhos de peneira foram conduzidos sob delineamento experimental de blocos incompletos, látice quadrado simples, parcialmente balanceado, sendo o experimento de peneira 5.5 um látice 6x6, com 34 progênies e duas testemunhas (BRASMAX BÔNUS IPRO - 8579 RSF e DM 81i84 RSF IPRO). O experimento de peneira 6.0 foi conduzido em látice 11x11, com 117 progênies e 4 testemunhas (BRASMAX BÔNUS IPRO - 8579 RSF, DM 81i84 RSF IPRO, M7739 IPRO e M8210 IPRO). O experimento de peneira 7.0 foi conduzido em látice 6x6 com 33 progênies e três testemunhas (BRASMAX BÔNUS IPRO - 8579 RSF, DM 81i84 RSF IPRO e M7739 IPRO).

Na safra 2022/23, foram avaliadas as progênies  $F_{2:4}$  em parcelas constituídas de duas linhas de três metros espaçadas em 0,6 metro, com três repetições, a densidade de semeadura foi ajustada conforme o grupo de maturidade absoluta das progênies. Os experimentos foram

conduzidos sob delineamento experimental de blocos incompletos, látice quadrado triplo parcialmente balanceado, o experimento de peneira 5.5 um látice 6x6, com 33 progênes e três testemunhas (BRASMAX BÔNUS IPRO - 8579 RSF, DM 81i84 RSF IPRO e M7739 IPRO). O experimento de peneira 6.0 foi conduzido em látice 6x6, com 34 progênes e 2 testemunhas (BRASMAX BÔNUS IPRO - 8579 RSF e DM 81i84 RSF IPRO), e o experimento de peneira 7.0 foi conduzido em látice 6x6 com 33 progênes e três testemunhas (BRASMAX BÔNUS IPRO - 8579 RSF, DM 81i84 RSF IPRO e M7739 IPRO).

A semeadura foi realizada manualmente, logo após foi realizada a aplicação de três doses de *Bradyrhizobium japonicum*, por meio de inoculante líquido, em jato dirigido nos sulcos de semeadura. Os demais tratos culturais foram realizados de acordo com a necessidade da cultura, empregando procedimento semelhante ao descrito por Soares *et al.* (2020), com o manejo de pragas, doenças e plantas daninhas.

### 3.4 Caracteres avaliados

Em ambas as safras as avaliações de campo, foram para os seguintes caracteres:

- Dias para o florescimento (DF): 50% das plantas da parcela, considerando todas as linhas apresentando florescimento pleno, estágio R2;
- Tipo de folíolo (TF): morfologia do trifólio das plantas, classificando progênes com folíolo lanceolado, ovalado ou segregante;
- Maturação absoluta (MA): 90% das plantas da parcela em estágio R8;
- Altura de plantas (AP): distância em centímetros do colo da planta até a extremidade da haste principal;
- Altura de inserção da primeira vagem (AV): distância em centímetros do colo da planta até o nó de inserção da primeira vagem;
- Índice de acamamento (IC): avaliado de acordo com Bernard *et al.* (1965), onde: 1, todas as plantas eretas; 2, algumas plantas inclinadas ou ligeiramente acamadas; 3, todas as plantas moderadamente inclinadas ou 25-50% das plantas acamadas; 4, todas as plantas severamente inclinadas ou 50-80% das plantas acamadas e 5, mais de 80% das plantas acamadas.

Para determinação dos componentes de produção, por ocasião da colheita, foram colhidas três plantas competitivas, isto é, em condições uniformes do experimento, de cada parcela, a fim de realizar as seguintes avaliações:

- Número de vagens por planta (NVP): contagem individual de vagens de cada planta;
  - Número de grãos por vagem (NGV): contagem do número de grãos de cada vagem.
- Com os dados de contagem, foi realizada uma regressão de segundo grau conforme modelo um e a estimativa do número de grãos por vagem médio que melhor representa a parcela, ou seja, o ponto máximo entre a distribuição dos dados de contagem, conforme modelo dois;

(1)

$$\bar{y} = \beta_0 + \beta_i x_i + \beta_j x_j^2 + \varepsilon_{ij}$$

(2)

$$\bar{x} = \frac{-\beta_i x_i}{2 * \beta_j x_j^2}$$

em que:

$\bar{y}$ : variável resposta prevista;

$\bar{x}$ : número de grãos por vagem que maximiza a função;

$\beta_0$ : Constante do modelo;

$\beta_i, \beta_j$ : coeficientes do modelo;

$x_i, x_j^2$ : termos polinomiais do modelo;

$\varepsilon_{ij}$ ; erro associado a variável resposta  $\bar{y}$ .

- Peso de cem grãos (PCG): mensuração da massa de cem grãos em gramas;
- Produtividade de grãos (PROD): determinada com base na colheita de todas as plantas da parcela. Após a padronização da umidade de grãos para 13% foi definida a produtividade, em  $\text{kg ha}^{-1}$ , a partir da conversão da área de cada parcela;

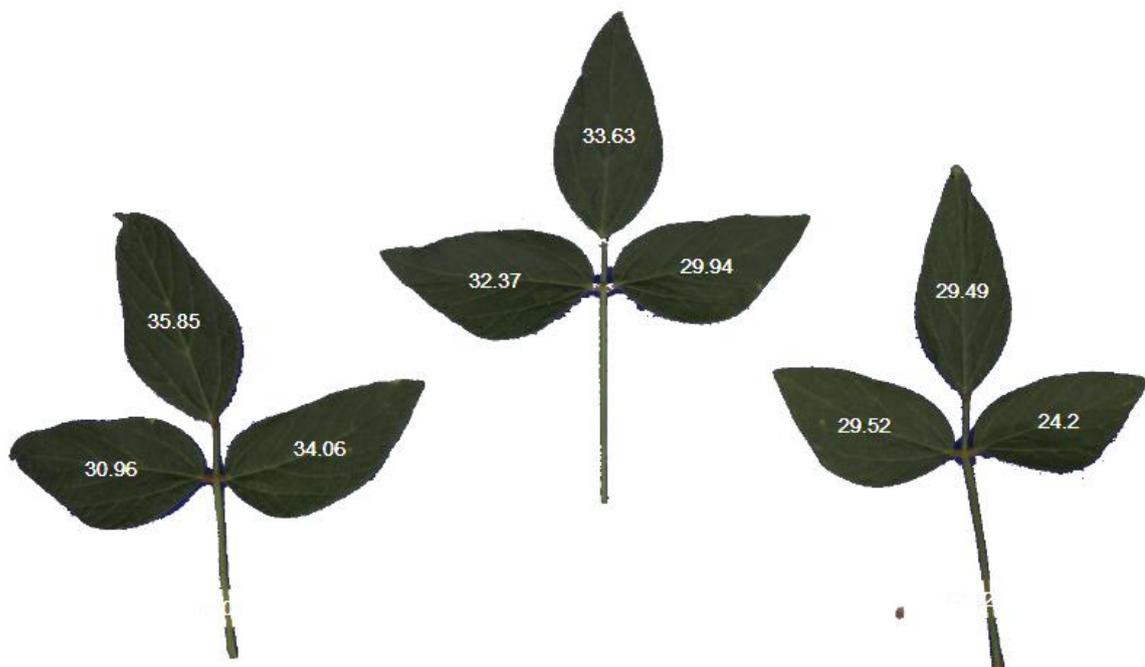
### 3.5 Uso de imagens na fenotipagem

Para processamento das imagens, foram coletados três trifólios no terço superior (estádio R4) por parcela nos três experimentos, peneira 5.5, 6.0 e 7.0 apenas os experimentos conduzidos em Lavras-MG na safra 2022/2023. A captura das imagens digitais RGB foi feita

usando uma câmera profissional modelo Canon EOS 60D D750, equipada com uma lente DX de 35mm, com motor de autofoco embutido, abertura de 1,8f e configuração ISO (sensor de captura de sensibilidade da câmera de luz) fixado em 1200 (Figura 2).

As imagens foram obtidas no formato de compressão “JPG”, sistema de cores RGB (vermelho, verde e azul). Para obter imagens de melhor qualidade, foi utilizada um estúdio, com luz artificial na parte superior e fundo que proporciona contraste com as folhas. Para extrair as informações, as imagens foram analisadas utilizando o pacote Pliman no ambiente R (OLIVOTO, 2021).

Figura 2 – Representação da avaliação de áreas dos folíolos mensurada nos experimentos de Lavras-MG na safra 2022/2023.



Fonte: Do autor (2023).

Os valores de área são representados nos números acima, em cm<sup>2</sup>

### 3.6.1 Análise dos dados

As análises estatísticas dos caracteres avaliados foram realizadas considerando a abordagem de modelos mistos seguindo o procedimento de melhor predição linear não viesada (BLUP - *Best linear unbiased prediction*) com a estimativa dos componentes de variância obtidos pelos métodos de máxima verossimilhança restrita (REML).

As variâncias genéticas entre as características foram testadas pelo teste de razão de verossimilhança (LRT), ao nível de 5% de probabilidade, considerando a aproximação de *qui-quadrado* na média de genótipos, de modo a identificar variação genética significativa para a

escolha de variáveis.

A análise individual dos dados foi realizada com o intuito de se testar a normalidade dos resíduos com base no teste de Shapiro-Wilk (SHAPIRO; WILK, 1965) e detectar a homogeneidade da variância entre as safras e locais conforme o teste  $F$  máximo (HARTLEY, 1950), utilizando o modelo três.

$$y_{ijk} = \mu + g_i + r_k + b_{j(k)} + \varepsilon_{ijk} \quad (3)$$

em que:

$y_{ijk}$ : observação do genótipo  $i$  no bloco  $j$  dentro da repetição  $k$ ;

$\mu$ : constante associada a todas as observações;

$g_i$ : efeito aleatório das genótipo  $i$ ;

$r_k$ : efeito aleatório da repetição  $k$ ;

$b_{j(k)}$ : efeito aleatório do bloco  $j$  dentro da repetição  $k$ ;

$\varepsilon_{ijk}$ : erro experimental aleatório associado a observação  $y_{ijk}$ .

Após os procedimentos básicos, a análise conjunta dos ambientes (local e safra) foi realizada a partir do modelo quatro.

$$y_{ijklm} = \mu + a_l + g_i + ga_{il} + r_{kl} + b_{j(kl)} + p_m + \varepsilon_{ijklm} \quad (4)$$

em que:

$y_{ijklm}$ : valor ajustado do genótipo  $i$  do bloco  $j$  dentro da repetição  $k$  alinhado com o ambiente  $l$ ;

$\mu$ : constante associada a todas as observações;

$a_l$ : efeito fixo do ambiente  $l$ ;

$g_i$ : efeito aleatório do genótipo  $i$ ;

$ga_{il}$ : efeito aleatório da interação genótipo  $i$  com o ambiente  $l$ ;

$r_{kl}$ : efeito fixo da repetição  $k$  alinhada ao ambiente  $l$ ;

$b_{j(kl)}$ : efeito aleatório do bloco  $j$  dentro da repetição  $k$  alinhado ao ambiente  $l$ ;

$p_m$ : efeito aleatório do tamanho de peneira  $m$ ;

$\varepsilon_{ijklm}$ : efeito aleatório do erro associado a observação  $y_{ijklm}$ .

A precisão experimental foi aferida por meio da estimativa do coeficiente de variação experimental (CVe) adotando a equação cinco e pela acurácia seletiva conforme equação seis (RESENDE; DUARTE, 2007).

(5)

$$CV = \frac{\sqrt{\sigma_E^2}}{\bar{X}}$$

em que:

$\sigma_E^2$ : Variância ambiental;

$\bar{X}$ : Média fenotípica geral.

(6)

$$r_{\hat{g}g} = \sqrt{1 - \left(\frac{PEV}{\sigma_G^2}\right)}$$

em que:

$PEV$ : Variância do erro da predição dos BLUP;

$\sigma_G^2$ : Variância genotípica.

A herdabilidade foi estimada a partir da fórmula sugerida por Cullis *et al.* (2006), como descrito no modelo sete.

(7)

$$Hc^2 = 1 - \frac{vBLUP}{2\sigma_G^2}$$

em que:

$Hc^2$ : Valor da herdabilidade para cada variável;

$vBLUP$ : Variância média da diferença entre dois BLUPs;

$\sigma_G^2$ : Variância genética.

Também foram estimadas as correlações genéticas e a fenotípica entre as variáveis analisadas, pelo modelo oito. No caso da correlação fenotípica, os valores foram testados com base no teste  $t$  ao nível de 0,05 de significância.

(8)

$$\rho_{xy} = \frac{\sigma_{xy}}{\sqrt{(\sigma_x^2 * \sigma_y^2)}}$$

em que:

$\rho_{xy}$ : correlação entre as variáveis  $x$  e  $y$ ;

$\sigma_{xy}$ : covariância entre as variáveis  $x$  e  $y$ ;

$\sigma_x^2$ : variância da variável  $x$ ;

$\sigma_y^2$ : variância da variável  $y$ ;

### 3.6.2 Resposta à seleção

Após a análise dos dados e estimativa dos parâmetros genéticos, o ganho esperado com a seleção direta foi calculado conforme o modelo nove, a partir do ranqueamento de progênies em função do ideótipo (FALCONER; MACKAY, 1996).

(9)

$$SD_k = i * p * \sigma_g * h$$

em que:

$SD_k$ : Ganho esperado com a Seleção Direta do caráter  $k$ ;

$i$ : intensidade de seleção, 15% (1,55);

$p$ : controle parental (adotado como igual a 1);

$\sigma_g$ : desvio padrão genético;

$h$ : raiz quadrada da herdabilidade da característica  $k$ .

Com o intuito de se aplicar seleção de forma indireta foi calculada a resposta correlacionada (ganho indireto), que corresponde ao ganho no caráter  $k$  em detrimento da seleção para o caráter  $y$ , conforme modelo dez (FALCONER; MACKAY, 1996).

(10)

$$RC_{k(y)} = i * p * r_{g_{ky}} * \sigma_{gk} * h_y$$

em que:

$RC_{k(y)}$ : resposta correlacionada em  $k$  pela seleção no caráter auxiliar  $y$ .

$i$ : intensidade de seleção padronizada, 15% (1,55);

$p$ : controle parental (será adotado como igual a 1);

$r_{g_{ky}}$ : correlação genética entre os caracteres  $k$  (principal) e  $y$  (auxiliar);

$\sigma_{gk}$ : desvio padrão genético de característica principal  $k$ ;

$h_y$ : raiz quadrada da herdabilidade da característica auxiliar  $y$ .

Para avaliar o mérito da seleção indireta em relação a seleção direta, o modelo 11

estima a razão das respostas correlacionada esperada e a resposta direta esperada. Para avaliar a viabilidade da realização da seleção indireta entre as variáveis. O testador do modelo 12 foi ajustado para avaliar a relação existente entre a razão das respostas correlacionada esperada e a resposta direta esperada (FALCONER; MACKAY, 1996).

(11)

$$R_{RC(SD)_k} = \frac{i_y r_{g_{ky}} h_y}{i_k h_k}$$

(12)

$$r_{g_{ky}} = r_{g_{ky}} * h_y$$

em que:

$R_{RC(SD)_k}$ : razão de mérito da resposta correlacionada do caráter em relação a seleção direta para a característica  $k$ ;

$i_y$ : intensidade de seleção padronizada para a característica  $y$  (auxiliar);

$i_k$ : intensidade de seleção padronizada para a característica  $k$  (principal);

$r_{g_{ky}}$ : correlação genética entre as características  $x$  e  $y$ ;

$h_y$ : raiz quadrada da herdabilidade da característica auxiliar  $y$ ;

$h_k$ : raiz quadrada da herdabilidade da característica principal  $k$ .

O ganho efetivo com a seleção direta em porcentagem ( $GS - \%$ ) foi calculado a partir das médias BLUPs da análise conjunta, pela equação 13.

(13)

$$GS - \% = \frac{GS}{\bar{X}_b} * 100$$

em que:

$GS - \%$ : ganho efetivo com a seleção direta para a característica  $i$ , em porcentagem;

$GS$ : diferencial de seleção com intensidade de 15%, sendo  $GS = \bar{X}_{bS} - \bar{X}_b$ ,

em que:

$\bar{X}_{bS}$ : média dos BLUPs das progênes selecionadas;

$\bar{X}_b$ : Média geral das predições BLUPs entre progênes.

### 3.6.3 Análise geométrica

A contribuição dos componentes de produção número de grãos por vagem, peso de

mil grãos e número de vagens por plantas para a produtividade foi mensurada usando a análise geométrica proposta por Grafius (1964) para as progênes selecionadas, conforme modelo 14. Para tal, os dados foram normalizados usando o critério Min-Max (JAIN; BHANDARE, 2013), conforme modelo 15.

(14)

$$N_{R_i} = \frac{n}{X_i + Y_i + Z_i}$$

em que:

$N_{R_i}$ : contribuição relativa da  $i$ -ésima variável ( $X_i, Y_i$  e  $Z_i$ );

$n$ : variável latente ( $X_i, Y_i$  e  $Z_i$ );

$X_i$ : número de vagens por planta da progênie  $i$ ;

$Y_i$ : número de grãos por vagens da progênie  $i$ ;

$Z_i$ : peso de mil grãos da progênie  $i$ .

(15)

$$X_{ijn} = \frac{X_{ij} - X_{minj}}{X_{maxj} - X_{minj}}$$

em que:

$X_{ijn}$ : Valor normalizado na escala de 0 a 1 da progênie  $i$  e variável  $j$ ;

$X_{ij}$ : Valor a ser transformado da progênie  $i$  e variável  $j$ ;

$X_{maxj}$ : Valor máximo do vetor da variável  $j$ ;

$X_{minj}$ : Valor mínimo do vetor da variável  $j$ .

Considerando que o ideótipo para este caso é um cubo com três dimensões iguais, a contribuição relativa de cada componente de produção é de 33,33%. Desta forma, foi estimada a soma de quadrados de desvios conforme modelo 16, de modo a estimar a discrepância entre a contribuição das variáveis analisadas para cada progênie selecionada.

(16)

$$SQI_i = (X_i - 0,333)^2 + (Y_i - 0,333)^2 + (Z_i - 0,333)^2$$

em que:

$SQI_i$ : soma de quadrados dos desvios em relação ao ideótipo para a progênie  $i$ .

### 3.6.4 Análise de agrupamento

Para melhor compreender a relação entre as variáveis estudadas neste trabalho, foi

empregada a análise multivariada de dissimilaridade entre as progênes avaliadas. Para tal, inicialmente, foi realizada a análise de trilha (OLIVOTO *et al.*, 2017) entre as variáveis estudadas de modo a realizar a dissimilaridade entre variáveis sem a presença de multicolinearidade. Como critério foi adotado valor inflacionário da variância (VIF) (DORMANN *et al.* 2012) inferior a dez. Desta forma, as variáveis utilizadas foram produtividade, número de grãos por vagem, peso de mil grãos, área de trifólios e tipo de folíolo.

Estas variáveis compõem um grupo de variáveis mistas (numéricas e categóricas). Desta forma, a estimativa da matriz de dissimilaridade entre as progênes foi realizada utilizando o método de Gower, conforme modelo 17, que além de um conjunto de variáveis mistas, contempla um conjunto de dados desbalanceados (FERREIRA, 2018).

(17)

$$d_{ij} = \frac{\sum_{n=1}^p d_{ij}^{(n)} \delta_{ij}^{(n)}}{\sum_{n=1}^p \delta_{ij}^{(n)}}$$

em que:

$d_{ij}^{(n)}$ : dissimilaridade entre os indivíduos  $i$  e  $j$ , para a  $n$ -enésima variável, sendo:

Quantitativas discretas:

$$d_{ij}^{(n)} = \begin{cases} 1 & \text{se } y_{in} \neq y_{jn} \\ 0 & \text{se } y_{in} = y_{jn} \end{cases}$$

Quantitativas contínuas:

$$d_{ij}^{(n)} = \frac{|y_{in} - y_{jn}|}{\max_n - \min_n}$$

em que:

$y_{in}$ : observação do indivíduo  $i$  para a  $n$ -enésima variável;

$y_{jn}$ : observação do indivíduo  $j$  para a  $n$ -enésima variável;

$\max_n$ : valor máximo do vetor da  $n$ -enésima variável;

$\min_n$ : valor mínimo do vetor da  $n$ -enésima variável;

$\delta_{ij}^{(n)}$ : variável indicadora de desbalanceamento: um para dados balanceados entre a indivíduos  $i$  e  $j$  e 0 para ao menos uma perdida. Se as duas variáveis forem perdidas, a observação é desconsiderada;

$p$ : vetor de variáveis analisadas.

De posse da matriz de dissimilaridade, foi então realizado o agrupamento hierárquico

da ligação média (UPGMA – *unweighted pair group method average*), com ponto de início na menor dissimilaridade, sendo este método escolhido em função da correlação cofrenética, que mede o quanto o agrupamento hierárquico explica a matriz de dissimilaridade original superior em relação aos demais métodos. Com o agrupamento, a matriz de dissimilaridade atualizada pelo modelo 18 foi utilizada para construir um dendrograma (FERREIRA, 2018).

(18)

$$\bar{d}_{(i,j)t} = \frac{d_{(i)t} + d_{(j)t}}{2}$$

em que:

$\bar{d}_{(i,j)t}$ : dissimilaridade média entre os indivíduos  $i$  e  $j$  em relação ao indivíduo  $t$  (ponto de fusão no dendrograma);

$d_{(i)t}$ : dissimilaridade do indivíduo  $i$  em relação ao indivíduo  $t$ ;

$d_{(j)t}$ : dissimilaridade do indivíduo  $j$  em relação ao indivíduo  $t$ .

Com o agrupamento hierárquico realizado, foi então estimado o ponto de corte para a formações de grupos, com base no método Mojena (1977), conforme modelo 19.

(19)

$$PCM = \frac{\sum_m \bar{d}_{(i,j)t}}{m} + \sigma_{\bar{d}_{(i,j)t}} * k$$

em que:

$PCM$ : ponto de corte Mojena;

$\sum_m \bar{d}_{(i,j)t}$ : somatório das dissimilaridades nos  $m$  pontos de fusão do dendrograma;

$\sigma_{\bar{d}_{(i,j)t}}$ : desvio padrão dos  $m$  pontos de fusão do dendrograma;

$k$ : constante, assumida como 2, conforme Mojena (1977).

### 3.6.5 Aspectos computacionais

O ambiente R (R Core Team, 2020) versão 4.1.3 foi utilizado para o tratamento dos dados e para o ajuste dos modelos utilizados. Utilizou-se o pacote sommer (COVARRUBIAS-PAZARAN, 2016) para o ajuste do modelo linear misto. As ferramentas disponíveis no pacote Tydiverse (WICKHAM *et al.*, 2019) foram utilizadas para a manipulação dos dados e para a elaboração dos gráficos apresentados.

#### 4 RESULTADOS E DISCUSSÃO

Inicialmente, destaca-se que os caracteres dias para o florescimento, dias para a maturação absoluta, altura das plantas, altura da inserção da primeira vagem e índice de acamamento das progênes nos anos agrícolas 2021/22 e 2022/23 em Lavras e Ijaci, foram utilizados apenas para a caracterização fenotípica e os resultados não serão apresentados neste presente trabalho. Neste sentido, no presente estudo, ênfase foi dada nos caracteres produtividade de grãos, componentes de produção e área dos trifólios, caracteres esses mais correlacionadas à produção.

As estimativas dos parâmetros genéticos e fenotípicos, para as características mencionados anteriormente, considerando a análise individual por ambiente foram apresentados na Tabela 1A do anexo. A precisão experimental, aferida pelas estimativas da acurácia seletiva e do coeficiente de variação experimental, variou em função dos ambientes e caracteres avaliados.

De acordo com Resende e Duarte (2007), a acurácia seletiva pode ser classificada como muito alta, alta, moderada e de baixa magnitude. A vantagem de se utilizar a acurácia para aferir a precisão experimental se perfaz no fato de que este parâmetro independe da estimativa da média do caráter em estudo. O menor valor de acurácia foi observado para a PROD que foi de 43%, magnitude esta considerada baixa. Para o NVP e a área de trifólios as acurácias foram classificadas como moderada, 61% e 69%, respectivamente. O caráter PMG e o NGV tiveram a acurácia seletiva classificada como de alta magnitude (Tabela 1). A acurácia e a herdabilidade são medidas intrinsecamente correlacionadas, isto é, espera-se baixa magnitude de acurácia em caracteres altamente influenciados pelo ambiente, como o caso da produtividade (FERREIRA, 2016).

O coeficiente de variação é considerado baixo quando seu valor é inferior a 10%; médio, quando a estimativa varia de 10% a 20%; alto, com os valores entre 20% e 30% e, muito alto, quando a magnitude for superior a 30% (PIMENTEL-GOMES, 2009). Os maiores CVes foram observados para os caracteres produtividade de grãos e NVP. Na safra 21/22, em Ijaci o CVe da produtividade foi de 37,42% e do NVP de 38,01%. Em Lavras os valores de CVe foram de 34,33% e 31,34% da produtividade e do NVP, respectivamente (Tabela 1A).

Segundo Pimentel Gomes (2009), o CV% depende do tamanho, forma da parcela e do caráter em consideração. Em caracteres poligênico, ou seja, quantitativo, espera-se valores de CV% altos, por ser altamente influenciado pelo ambiente. Esse resultado pode ser

explicado pelo fato que na geração F<sub>2:3</sub> a parcela experimental foi composta por apenas uma linha de um metro. Este cenário é comum nas fases iniciais dos programas de melhoramento devido ao fato das sementes das progênies selecionadas serem oriundas de uma única planta, e por consonância ter-se uma limitação de sementes. Contudo, deve-se comentar que no presente trabalho optou-se por adotar uma parcela experimental menor (0,6 m<sup>2</sup>), porém, com a condução dos experimentos em dois locais. No que tange o tamanho da parcela experimental, Martin *et al.* (2005), afirmam que o tamanho ótimo para experimentos com cultivares de soja é de 3,96 m<sup>2</sup>, por sua vez, González (2013), afirma que esse valor é de 5,94 m<sup>2</sup>.

Normalmente, nos programas de melhoramento com o objetivo de se obter estimativas mais precisas e acuradas, realiza-se a análise conjunta. Neste âmbito, as estimativas dos parâmetros genéticos e fenotípicos considerando os quatro ambientes em conjunto estão apresentados na Tabela 1. O maior CVE foi estimado em 34,44% para a produtividade. Ribeiro *et al.* (2020), estudando progênies de soja oriundas de um programa de seleção recorrente obtiveram estimativas do CVE para a produtividade de 22,69%. Contudo, deve-se destacar que no trabalho de Ribeiro *et al.* (2020) foram analisadas conjuntamente três gerações de endogamia testadas em mais anos agrícolas e mais locais o que propicia um menor CVE. Embora o coeficiente de variação seja utilizado como uma medida para avaliar o comportamento do conjunto de dados, esse normalmente não é utilizado de forma isolada e sim em conjunto com outras estimativas.

O teste da razão de verossimilhança evidenciou que para todas as características e ambientes estudados a variância genética foi significativa, com exceção para o NVP em Ijaci na safra 21/22, denotando assim a existência de variabilidade genética entre as progênies, fato este preponderante para a realização do presente estudo. Reforçando ainda os comentários anteriores, destaca-se que a herdabilidade para produtividade de grãos na safra 22/23 foi maior do que 40%, assim como, para o PMG em todos os ambientes, este fato destaca que existe possibilidade de sucesso com a seleção para este caráter dentre as progênies avaliadas (Tabela 1A).

Para o teste da razão de verossimilhança evidenciou que para todas as características estudadas a variância genética foi significativa. Uma premissa básica da genética e melhoramento de plantas é que a efetividade da seleção depende da existência de variabilidade genética. A magnitude dessa variabilidade genética, por sua vez, pode ser estimada pela herdabilidade que se refere à proporção da variação fenotípica que é devida ao componente genético (RAMALHO *et al.*, 2012).

A herdabilidade estimada para produtividade foi de 18%. Para as demais características as estimativas foram superiores a 38%. Silva *et al.* (2022), estudando os parâmetros fenotípicos e genotípicos de soja obtiveram valores de 50,14% de herdabilidade para a produtividade de grãos, magnitude esta superior a obtida no presente estudo. Contudo, se faz necessário ressaltar que esses resultados foram obtidos em experimentos com gerações mais avançadas de endogamia o que propicia maior estabilidade entre os genótipos e por consonância, maior precisão nas avaliações. A baixa magnitude da herdabilidade para a produtividade no presente trabalho pode ser explicada também pelo erro associado as estimativas. Na variável produtividade a variância do erro experimental superou a variância genética em quase três vezes mais (Tabela 1).

DESTRO *et al.* (1987) e BACKES *et al.* (2002) constataram que o caráter produtividade de grãos sofre maior influência ambiental, resultando normalmente em menores valores de herdabilidade em comparação aos demais caracteres agronômicos estudados. Além disso, para o mesmo caráter, destaca-se a variância devido à dominância, que também colabora para a obtenção de baixas estimativas de herdabilidade no sentido restrito.

Tabela 1 – Estimativa de parâmetros genéticos de progênie de soja em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG.

PARÂMETROS	PROD	PMG	NGV	NVP	AT
QME	2.419.456,00	68,58	0,03	1832,224	186,83
$\sigma_g^2$	93.140,52	200,53	0,02	250,34	159,81
$h_a^2$	0,18	0,52	0,75	0,38	0,48
$r_{ap}$	0,43	0,72	0,86	0,61	0,69
CVe	34,44	5,10	7,51	34,02	18,37
LRT	0,04763*	4,32e <sup>-87</sup> ***	8,69e <sup>-35</sup> ***	2e-0,5***	1,08e <sup>-76</sup> ***

Média geral ( $\mu$ ), quadrado médio do erro (QME), variância genética ( $\sigma_g^2$ ), herdabilidade da média no sentido amplo ( $h_a^2$ ), acurácia seletiva ( $r_{ap}$ ), coeficiente de variação experimental (CVe) e teste da razão de verossimilhança (LRT). Produtividade de grãos (kg ha<sup>-1</sup>) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta (NVP), área dos trifólios (cm<sup>2</sup>) (AT).

\*\*Significativo ao nível de 0,05 de probabilidade, considerando a aproximação de qui-quadrado.

\*\*\*Significativo ao nível de 0,01 de probabilidade, considerando a aproximação de qui-quadrado.

Fonte: Do autor (2023).

Os valores máximos, mínimos, médios e a amplitude de variação das médias BLUP conjuntas, para todos os caracteres avaliados estão apresentados na Tabela 2. As estimativas das médias BLUP evidenciam a grande variação entre os genótipos para os caracteres

avaliados e reforçam os comentários apresentados anteriormente. A produtividade de grãos variou de 4833,29 kg ha<sup>-1</sup> da cultivar BMX Bônus a 4093,87 kg ha<sup>-1</sup> da progênie 167. Uma amplitude expressiva de 739 kg. Muito embora a testemunha BMX Bônus foi o tratamento de maior produtividade de grãos, 24,45% das progênies apresentaram produtividade superior a 4500 kg ha<sup>-1</sup>, denotando assim o potencial produtivo destas e a possibilidade de sucesso com a seleção.

A maior amplitude de variação foi observada no caráter área dos trifólios (Tabela 2). O genótipo de maior área foliar obteve magnitude duas vezes maior do que o de menor área, evidenciando a grande variabilidade observada para essa característica. Para o NVP observou-se também grande variabilidade com uma amplitude observada de 62,64. Resultados similares foram obtidos por Carmo *et al.* (2018) que estudando o desempenho agrônomico de cultivares de soja em oito ambientes diferentes observaram uma amplitude de 21,70 para a característica NVP. A alta magnitude para estas características pode ser explicada através do fenótipo dos genitores que são contrastantes para tais características. Uma menor variação fenotípica foi observada no caráter NGV quando comparado com os demais caracteres (Tabela 2).

Tabela 2 – Estimativa de parâmetros estatísticos de progênies de soja em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG.

<b>PARÂMETROS</b>	<b>PROD</b>	<b>PMG</b>	<b>NGV</b>	<b>NVP</b>	<b>AT</b>
Máximo	4.833,29	200,81	2,77	164,04	106,56
Mínimo	4.093,87	133,16	2,22	101,40	53,42
Média	4.442,78	161,47	2,55	126,27	74,02
Amplitude	739,42	67,65	0,54	62,64	53,14

Produtividade de grãos (kg ha<sup>-1</sup>) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta (NVP), área dos trifólios (cm<sup>2</sup>) (AT).

Fonte: Do autor (2023).

Moro *et al.* (2021), estudando seis diferentes densidades de semeaduras observaram que o NGV teve baixa variação nos diferentes ambientes. No mesmo estudo os autores observaram uma amplitude de apenas 1,8 gramas para PMG nos diferentes ambientes estudados, não corroborando os achados obtidos no presente estudo, no qual observou-se para o PMG uma amplitude de 67,65 gramas. Essas diferenças observadas podem ser explicadas devido a fatores genéticos, por exemplo, uma vez que os genótipos são distintos, as amplitudes maiores são observadas principalmente em cruzamentos contrastantes.

O ganho com a seleção é o principal parâmetro a ser analisado em um programa de melhoramento. A seleção direta visa a seleção das melhores progênes considerando um único caráter. Neste sentido, verifica-se na Tabela 3 as estimativas do ganho esperado com a seleção de progênes de soja, considerando uma intensidade de seleção de 15%, valor médio encontrado na literatura (RAMALHO et al., 2012; SOARES et al., 2020).

A maior porcentagem do ganho com a seleção direta foi observada na característica área dos trifólios. A seleção direta para NVP propiciou um ganho esperado com a seleção de 12,12%. Por outro lado, para PMG observou-se um ganho esperado de 10,34%, magnitude esta superior a observada para o NGV e a produtividade de grãos (Tabela 3).

Tabela 3 – Ganhos com a seleção direta para a seleção de indivíduos superiores a uma intensidade de 15% em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG.

PARÂMETROS	PROD	PMG	NGV	NVP	AT
Média das progênes	4.439,19	161,30	2,56	126,36	73,74
Média P.S.	4.645,24	177,98	2,71	141,68	88,14
$h_a^2$	0,18	0,52	0,75	0,38	0,48
GS	206,05	16,68	0,15	15,32	14,40
GS - %	4,64	10,34	5,85	12,12	19,52

Média das progênes selecionados por variável (Média P.S.), herdabilidade no sentido amplo ( $h_a^2$ ), ganho com a seleção com relação à média das progênes (GS), porcentagem do ganho com a seleção com relação à média das progênes (GS - %). Produtividade de grãos (kg ha<sup>-1</sup>) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta (NVP), área dos trifólios (cm<sup>2</sup>) (AT).

Fonte: Do autor (2023).

Soares *et al.* (2020), estudando os parâmetros genéticos e fenotípicos de progênes de soja, em experimentos conduzidos na mesma região na qual o presente estudo foi realizado, obtiveram com a seleção direta para a característica produtividade um ganho de seleção de 9,04%, em uma proporção de seleção de 15%. Contudo, é oportuno ressaltar que neste estudo os autores utilizaram progênes oriundas do programa de seleção recorrente para produtividade de grãos, técnica que possibilita explorar uma maior variabilidade genética. Além disso, foi utilizada uma parcela experimental maior e mais locais, o que propiciou maior precisão, corroborando assim em maior magnitude do GS%.

A Figura 3 apresenta um ranking com os 28 genótipos de maior produtividade, e a contribuição de cada componente na produtividade de grãos. A porcentagem de contribuição

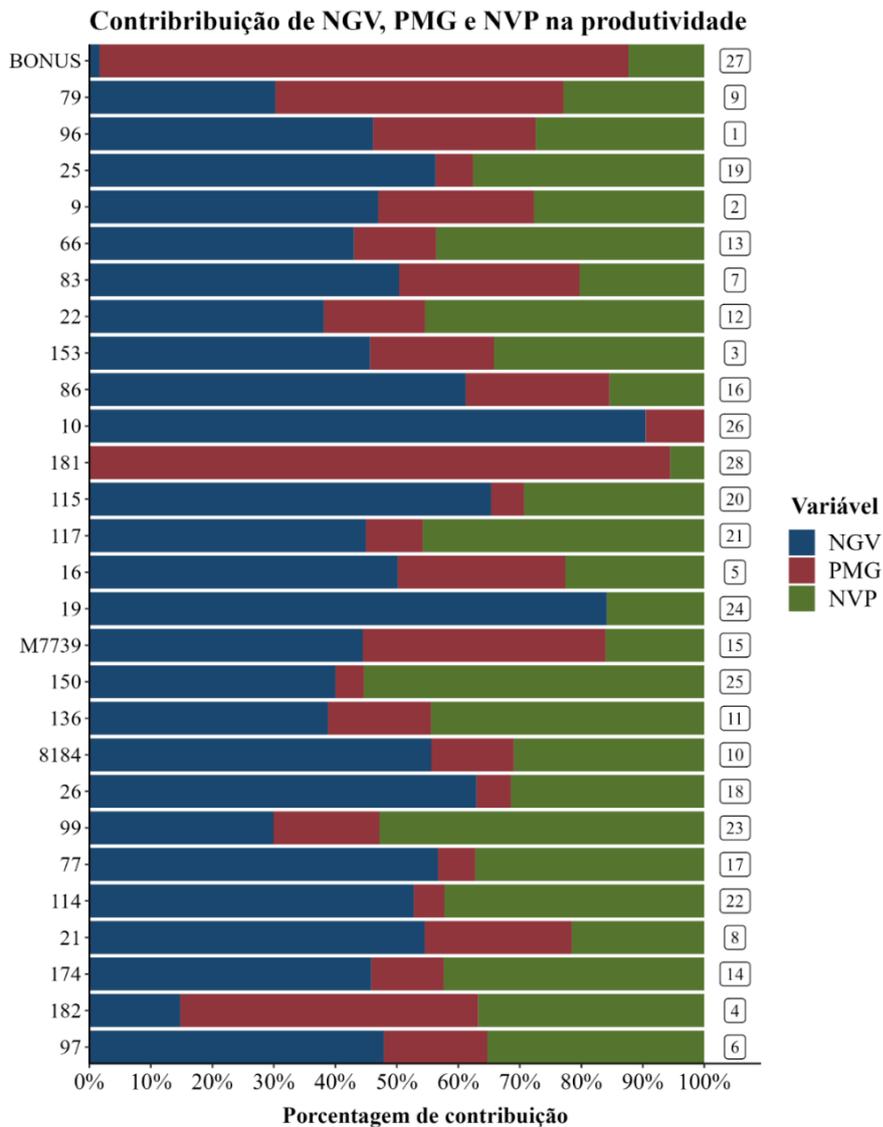
dos componentes variou em função do genótipo. Em 18 desses tratamentos o componente de maior relevância para a produtividade foi o NGV. Por sua vez, o NVP teve maior porcentagem de contribuição em seis tratamentos. O PMG por outro lado, predominou em apenas quatro tratamentos. Navarro Júnior e Costa (2002) encontraram resultados similares ao nosso estudo, evidenciando que a contribuição dos componentes de produção para a produtividade de grãos varia em função do genótipo estudado.

Isso significa que as proporções em que os diferentes fatores influenciam a produtividade não são as mesmas para todos os genótipos de soja estudados. Essa variação nos efeitos dos componentes pode ser resultado das características genéticas específicas de cada progênie, assim como da interação genótipo por ambiente, permitindo assim, compreender como as características das diferentes cultivares de soja podem afetar sua produtividade. A interação genótipo x ambiente pode ser reduzida, utilizando-se cultivares específicas para cada ambiente, ou utilizando-se cultivares com ampla adaptabilidade e boa estabilidade, ou estratificando-se a região considerada em sub-regiões com características ambientais semelhantes, dentro das quais a interação passa a ser não-significativa (Ramalho et al., 1993).

Segundo a teoria de Grafius (1964) o genótipo o ideal seria aquele que tivesse a contribuição equalitária dos componentes de produção na produtividade de grãos, ou seja, 33,33% dividido para os três caracteres estudados. O genótipo que mais se aproximou do ideótipo de Grafius foi a progênie 96, com 46,12% de contribuição para NGV, 27,44% para o NVP e 26,44% para PMG (Figura 3). Contudo, ao contrário do que Grafius esperava, a progênie 96 não foi o genótipo de maior produtividade, mas sim a cultivar BMX BÔNUS, que no ranking de equivalência da contribuição dos componentes ficou apenas em penúltimo lugar, denotando que nem sempre o genótipo de maior equivalência entre os componentes é o de maior produtividade.

Isso significa que diferentes características das plantas podem ter um impacto variável na produção final de grãos. Esses resultados corroboram com os relatos por Lima e Ramalho, (2016) com a cultura do feijoeiro, que estudando a viabilidade da aplicação da geometria no melhoramento do feijoeiro concluíram que o emprego dessa metodologia no melhoramento de plantas é de baixa eficiência.

Figura 3 – Ranqueadas da contribuição dos caracteres NGV, PMG e NVP na produtividade de grãos de progênies de soja em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois ambientes Lavras-MG e Ijaci-MG. Peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta (NVP).



Os números a direita referem-se ao ranking de classificação das progênies conforme a contribuição dos componentes NGV, PMG e NVP, seguindo o ideótipo de Grafius.

Fonte: Do autor (2023).

A seleção de uma característica pode influenciar a resposta em outra, favoravelmente ou negativamente ao ideótipo desejado. Isso ocorre em função da magnitude e direção da correlação genética observada entre um par de variáveis (RODRIGUES *et al.*, 2010). Desta forma, a resposta correlacionada, é um parâmetro importante a ser mensurado para a seleção indireta, principalmente para características de baixa herdabilidade como é o caso da produtividade (RIBEIRO *et al.*, 2020; MARTINS, 2020; SOARES *et al.*, 2020). A análise de

trilha evidenciou alta colinearidade, entre as características NGV e NVP, ou seja, demonstram uma relação linear (estão altamente relacionadas) assim, optou-se por excluir o caráter NVP das análises de resposta correlacionada e do agrupamento de genótipos (Tabela 4 e Figura 4 e 5).

A maior porcentagem de ganho indireto para produtividade foi considerando a seleção para número de grãos por vagem, 1,07% (Tabela 4), complementando os resultados descritos na Figura 3. A seleção indireta para as características PMG e AT resultou em ganhos de 0,12% e 0,46%, respectivamente. Ao seleciona-se diretamente para a produtividade tem-se ganhos positivos para as três características, sendo o maior ganho para a área dos trifólios que foi de 2,81%, seguido pelo NGV com ganho de 1,28% e PMG com 0,18%. Observou-se também uma correlação inversamente proporcional entre o número de grãos por vagem e o PMG de tal maneira que a seleção indireta para PMG resulta em -7,41% no número de grãos por vagem e a seleção para o NGV resulta em -9,36% no PMG, essa foi a maior magnitude entre as correlações estudadas (Tabela 4).

Tabela 4 - Resposta correlacionada e razão genética para os parâmetros avaliados em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG sobre uma intensidade de seleção de 15%.

<b>PARÂMETROS</b>	<b>PROD</b>	<b>PMG</b>	<b>NGV</b>	<b>AT</b>
$h_x$	0,42	0,72	0,86	0,69
SD (%)	4,64	10,34	5,85	19,52
RC. PROD (%)	-	0,18	1,28	2,81
$R_{RC(SD)}$	-	0,01	0,08	0,05
$r_{g_{ky}} h_y$	-	0,01	0,07	0,04
RC. PMG (%)	0,12	-	-7,41	-4,98
$R_{RC(SD)}$	0,04	-	-0,51	-0,10
$r_{g_{ky}} h_y$	0,01	-	-0,44	-0,07
RC. NGV (%)	1,07	-9,36	-	-1,27
$R_{RC(SD)}$	0,36	-0,74	-	-0,02
$r_{g_{ky}} h_y$	0,15	-0,53	-	-0,01
RC. AT (%)	0,46	-1,22	-0,24	-
$R_{RC(SD)}$	0,15	-0,09	-0,01	-
$r_{g_{ky}} h_y$	0,06	-0,07	-0,01	-

Raiz quadrada da herdabilidade média dos tratamentos no sentido amplo ( $h_x$ ), Ganho com a seleção direta (SD), resposta correlacionada para a  $n$  (coluna) característica com a seleção indireta realizada para a variável  $m$  (linha) (RC. $n$ ), razão entre a resposta correlacionada e o ganho com a seleção direta ( $R_{RC(SD)}$ ) e, em itálico, testador para verificar a eficiência da seleção indireta sobre a seleção direta ( $r_{g_{ky}} h_y$ ). Produtividade de grãos (kg ha<sup>-1</sup>) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), área dos trifólios (cm<sup>2</sup>) (AT).

Fonte: Do autor (2023).

Públio Júnior *et al.* (2018) concluíram que o aumento na produtividade de grãos poderá ser obtido através da seleção indireta das características: número de vagens por planta, peso da vagem e número de grãos por vagem em genótipos de feijão. Fato este não observado no presente estudo. Apesar da correlação positiva entre os caracteres destaca-se que a magnitude dessa correlação não foi significativa, o que denota que a seleção indireta para PMG, NGV e AT poderá não ser efetiva para resultar ganhos na produtividade.

A adoção da seleção indireta depende da estimativa da razão entre a resposta correlacionada e o ganho com a seleção direta ( $r_{g_{ky}} h_y$ ). Este parâmetro é significativo quando  $r_{g_{ky}} h_y$  for superior a estimativa da raiz quadrada da herdabilidade do caráter principal. Analisando a raiz quadrada da herdabilidade para a produtividade ( $h_x$ ) que foi de 0,42, verificou-se que, nenhum dos valores estimados de  $r_{g_{ky}} h_y$  foram superiores a  $h_x$  da produtividade evidenciando então que o emprego desta estratégia não é eficiente. Esses resultados apoiam os encontrados por Wei e Molin (2020), que em seus estudos com estimadores da produtividade de soja concluíram que a produtividade da soja apresenta uma baixa relação linear com o peso de mil grãos.

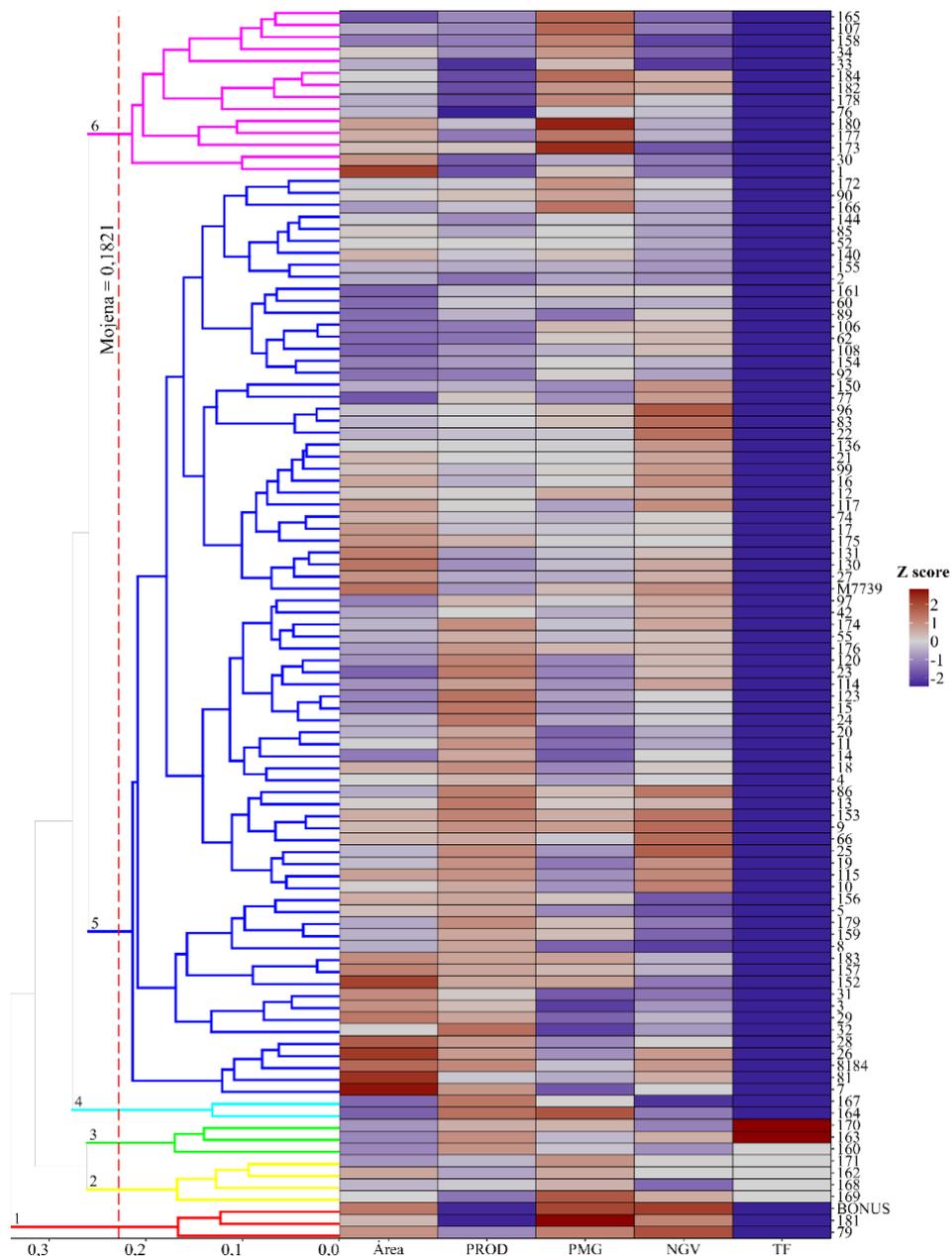
Uma possível explicação para estes resultados é a complexa interação entre os diferentes componentes de produção, ou seja, esses caracteres podem ter uma relação inversa ou não linear entre si, e a seleção considerando apenas um deles pode não refletir adequadamente a produtividade geral. Assim, este estudo sugere que a abordagem mais eficiente para selecionar linhagens superiores de soja é focar na seleção direta para produtividade de grãos. Essa estratégia contempla a interação complexa que existe entre os diferentes fatores que afetam a produtividade, incluindo o número de plantas por área, o número de vagens por planta, o número de grãos por vagem, o tamanho dos grãos e o peso médio dos grãos.

O peso de mil grãos, rendimento de peneira, e o número de grãos por vagem são características inversamente proporcionais. Essa relação indica que, quando os grãos são maiores, isto é, possuem maior peso individual, o número de grãos por vagem tende a ser menor, e vice-versa. Assim, um aumento no tamanho dos grãos pode levar a uma diminuição

no número total de grãos formados por planta. Logo, ao considerar todos esses fatores em conjunto, a seleção direta considerando a produtividade de grãos permite uma avaliação mais abrangente e precisa dos genótipos superiores, identificando aqueles que apresentam melhor desempenho global em termos de produção de grãos. Dessa forma, essa estratégia tem maior potencial para aumentar efetivamente a produtividade da cultura da soja, proporcionando melhores resultados aos agricultores e contribuindo para o avanço da sojicultura.

O dendrograma considerando a área de trifólios, produtividade de grãos, PMG, NGV e o tipo de folíolo apresentado na Figura 4 denota que os tratamentos foram divididos em seis grupos distintos.

Figura 4 - Dendrograma de progênies de soja em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois ambientes Lavras-MG e Ijaci-MG.



Para uma dada característica, as cores descrevem o quão próximo é uma progênie em relação à outra, em que valores positivos na escala de cores representam menor similaridade entre as progênies, e os negativos, maior proximidade.

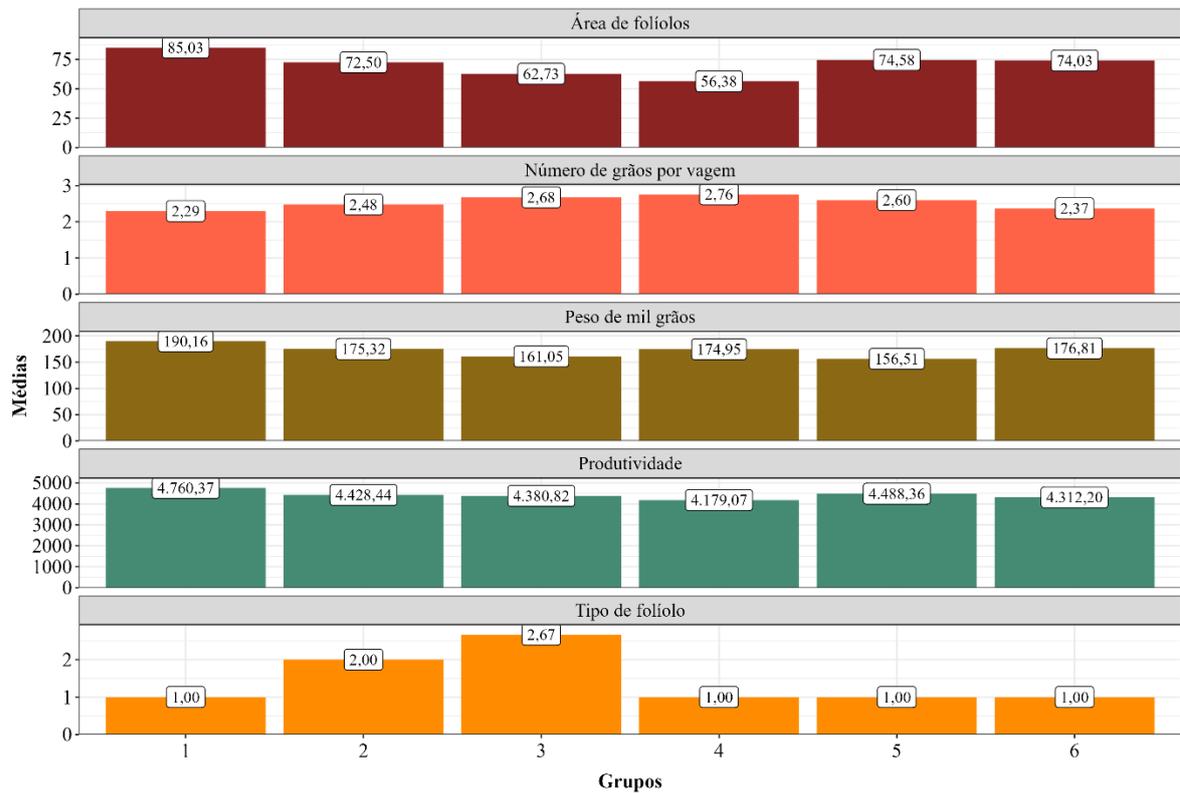
Fonte: Do autor (2023).

Para a obtenção dos grupos adotou-se o método de mojena (1977), no qual estima o tamanho relativo dos níveis de fusões, determinando-se uma distância máxima para que as progênies compreendam à um mesmo grupo. Verifica-se que o grupo 4 é o mais restrito, ou seja, com o menor número de progênies similares. Para as características avaliadas, neste grupo, os genótipos (progênies 167 e 164) têm maior semelhança para o caráter TF, mas diferem para PMG, convergindo na correlação entre o tipo de trifólio e número de grãos (mesmo trifólio, semelhança no número de grãos, conseqüentemente, no PMG). O TF por ser uma característica discreta apresenta menor variação genotípica, em função da sua herança monogênica, com progênies assumindo fenótipos ovalados, lanceolados ou segregantes. Para produtividade, uma característica contínua, poligênica, observa-se grande variação fenotípicas entre as progênies.

Como observado na Figura 5, o grupo de maior produtividade foi o grupo 1 com média de 4760,37 kg ha<sup>-1</sup>. Este grupo foi formado por três tratamentos, a testemunha BMX Bônus e duas progênies a 181 e a 79. Observou-se também a maior área dos trifólios com 85,03 cm<sup>2</sup> em média, e o maior PMG com 190,16 gramas na média, isto é, quanto maior a área de trifólios das plantas de soja, maior tende a ser a produtividade. Essa relação é compreensível, já que a área foliar maior permite maior captação de luz solar e, conseqüentemente, maior capacidade fotossintética da planta, o que pode resultar em uma produção de grãos mais abundante.

O grupo de menor produtividade foi o grupo 4 com 4179,07 kg ha<sup>-1</sup> (Figura 5). Neste grupo observou-se também a menor área de trifólios 56,38 cm<sup>2</sup>, bem como o maior número de grãos por vagem 2,76. Em trabalho apresentado por Sayama *et al.* (2017), estudando a relação entre o número de grãos por vagem e o tipo de folíolo em soja, os autores constataram que o gene *Ln* da soja controla de forma pleiotrópica as características número de grãos por vagem e tipo de folíolo. Corroborando os resultados obtidos no presente estudo, isto é, um maior NGV esteve associado ao fenótipo de folíolo lanceolado.

Figura 5 - Grupos formados por agrupamento de progênies de soja em análise conjunta, de quatro ambientes, geração  $F_{2:3}$  na safra 21/22 e  $F_{2:4}$  na safra 22/23 em dois ambientes Lavras-MG e Ijaci-MG.



Tipo de folíolo, ovalado (1), lanceolado (2), segregando (3).

Fonte: Do autor (2023).

## **5 CONCLUSÕES**

A seleção indireta, considerando os caracteres número de grãos por vagem ou peso de mil grãos, não figura-se a melhor estratégia para aumentar a produtividade de grãos na cultura da soja, devido à complexa interação entre os componentes de produção.

A melhor abordagem para selecionar linhagens superiores de soja é focar diretamente na produtividade de grãos, pois esta estratégia contempla a interação complexa que existe entre os diferentes fatores que afetam a produtividade e, por consonância, permite a identificação de genótipos superiores mais eficazmente.

## REFERÊNCIAS

**Abiove - Site.** Disponível em: <<https://abiove.org.br/estatisticas/exportacoes-do-complexo-soja/>>. Acesso em: 16 set. 2023.

ACKES, R.L.; REIS, M.S.; SEDIYAMA, T.; CRUZ, C.D.; TEIXEIRA, R. C. Estimativas de parâmetros genéticos em populações F5 e F6 de soja. *Revista Ceres*, Viçosa, v.49, n.282, p.201-216, 2002.

ARANTES, N. E.; SOUZA, P. I. de M. de. **Cultura da soja nos cerrados.** Piracicaba: Potafós, p. 267-268, 1993.

ARAUS, J. L. *et al.* **Translating high-throughput phenotyping into genetic gain.** *Trends in plant science*, v. 23, n. 5, p. 451-466, 2018. doi: 10.1016/j.tplants.2018.02.001

ARAUS, J.L., BUCHAILLOT, M.L., KEFAUVER, S.C. **High Throughput Field Phenotyping.** In: REYNOLDS, M. P., BRAUN, H. J. *Wheat Improvement.* Springer, Cham. 2022. 629 p.

BERNARD R. L.; CHAMBERLAIN, D. W.; LAWRENCE, R. E. **Results of the cooperative uniform soybean tests Part I.** *Uniform Soybean Tests, Northern Region*, 1965. 134 p.

BERNARDO, R. **Breeding for quantitative traits in plants.** 2 Ed., Woodbury, Minnesota: Editora Stemma Press, 2010. 400 p.

BOREM, A. (Org.). **Hibridação Artificial de Plantas.** 2a. ed. Viçosa, MG: Editora UFV, v.1. 2009. 625 p.

CARGNIN, A. et al. Interação entre genótipos e ambientes e implicações em ganhos com seleção em trigo. **Pesquisa Agropecuária Brasileira**, v. 41, n. 6, p. 987–993, 2006.

CARMO, E. L. *et al.* **Desempenho agrônomico da soja cultivada em diferentes épocas e distribuição de plantas.** *Revista de Ciências Agroveterinárias*, v. 17, n. 1, p. 61-69, 2018. doi: 10.5965/22381171171201806

CARNEIRO, A. K. et al. Stability analysis of pure lines and a multiline of soybean in different locations. **Crop Breeding and Applied Biotechnology**, v. 19, n. 4, p. 395–401, 2019.

CARNEIRO, V. Q. **Aplicativos computacionais para o melhoramento genético fundamentados em análise de imagens e inteligência computacional.** 2018.

COELHO, A. D. F. *et al.* **Herdabilidade e correlações da produção do feijão e dos seus componentes primários, nas épocas de cultivo da primavera-verão e do verão-outono.** *Ciência Rural*, v. 32, n. 2, p. 211-216, 2002. doi: 10.1590/S0103-84782002000200005

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da Safra Brasileira de Grãos**, v. 10, safra 2022/23, n. 9, nono levantamento, julho 2023. Brasília, 2023.

COVARRUBIAS-PAZARAN, G. **Genome Assisted Prediction of Quantitative Traits**

**Using the R Package sommer.** PLoS ONE, v. 11, n. 6, p. e0156744. 2016.  
doi:10.1371/journal.pone.0156744.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. **Modelos biométricos aplicados ao melhoramento genético.** v. 1, 2012. 514 p.

CULLIS, B. R.; SMITH, A. B.; COOMBES, N. E. **On the design of early Generation variety trials with correlated data.** Journal of agricultural, biological, and environmental statistics, v. 11, n. 4, p. 381-393, 2006. doi: 10.1198/108571106X154443

DALLASTRA, A. *et al.* **Multivariate approach in the selection of superior soybean progeny which carry the RR gene.** Revista Ciência Agronômica, v. 45, n. 3, p. 588-597, 2014.

DESTRO, D.; SEDIYAMA, T.; SILVA, J.C.; SEDIYAMA, C.S.; THIÉBAUT, J.T.L. Estimativas de herdabilidade de alguns caracteres, em dois cruzamentos de soja. Pesquisa Agropecuária Brasileira, Brasília, v.22, p.291-304, 1987.

DORMANN, C. F. *et al.* **Collinearity: a review of methods to deal with it and a simulation study evaluating their performance.** Ecography, v. 36, p. 27-46, 2013. doi: 10.1111/j.1600-0587.2012.07348.x

ELMERICH, C. *et al.* Identification of Eco-climatic factors driving yields and Genotype by Environment Interactions for yield in early maturity soybean using crop simulation. **Agronomy (Basel, Switzerland)**, v. 13, n. 2, p. 322, 2023.  
em duas épocas de semeadura. Bioscience Journal, Uberlândia, v. 28, n. 6, p.877-888,

FALCONER, D. S. **Introdução à genética quantitativa.** Viçosa-MG. Tradução de: Martinho de Almeida e Silva e José Carlos Silva.: UFV. 1987.

FALCONER, D. S.; MACKAY, T. F. **Introduction to quantitative genetics.** London: Longman. 1996.

FARIAS, J. R. B.; NEPOMUCENO, A. L.; NEUMAIER, N. **Ecofisiologia da soja.** 2007.

FERREIRA, D. F. **Estatística multivariada 3ª edição.** Editora UFLA. Lavras – MG. 2018.

GONZÁLEZ, G. G. H. **Estimação do tamanho ótimo de parcelas com aplicação na cultura da soja.** 2013. 85 f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, 2013.

GRAFIUS, J. E. **A geometry for plant breeding.** Crop Science, v. 4, n. 3, p. 241-246, 1964. doi: 10.2135/cropsci1964.0011183X000400030001x

HERMAN, T. K. *et al.* Evaluation of wild perennial *Glycine* species for resistance to soybean cyst nematode and soybean rust. **Plant breeding**, v. 139, n. 5, p. 923–931, 2020.

HOWARD, R.; JARQUIN, D. Predição genômica usando imagem de cobertura de dossel e informações genotípicas em soja por meio de modelo híbrido. *Bioinformática evolutiva online*

, v. 15, p. 1176934319840026, 2019.

KALAIVANI, R.; MURUGANAND, S.; PERIASAMY, A. **Identifying the quality of tomatoes in image processing using matlab.** International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, v. 2, n. 8, p. 3525-3531, 2013.

LANA, A. M. Q.; CARDOSO, A. A.; CRUZ, C. D. **Herdabilidades e correlações entre caracteres de linhagens de feijão obtidas em monocultivo e em consórcio com milho.** Ciência Rural, Santa Maria, v. 33, n. 6, p. 1031-1037, 2003. doi:10.1590/S0103-84782003000600006

LANGENBACH, C. et al. Fighting Asian soybean rust. **Frontiers in plant science**, v. 7, 2016.

LEITE, W. S. et al. **Estimativas de parâmetros genéticos e correlações entre caracteres agronômicos em genótipos de soja.** Nativa, v. 3, n. 4, p. 241–245, 2015. doi: 10.14583/2318-7670.v03n04a03

LI, Y. *et al.* **High-throughput phenotyping analysis of maize at the seedling stage using end-to-end segmentation network.** PLoS ONE, v. 16, n. 1, p. 1-19. 2021. doi: 10.1371/journal.pone.0241528

LIMA, J. G.; RAMALHO, M. A. P. **Geometry applied to breeding common beans (*Phaseolus vulgaris*).** Genetics and Molecular Research, v. 15, n. 2, p. gmr.15027873, 2016. doi: 10.4238/gmr.15027873

MAIMAITIJIANG, M. et al. Modelo de volume ponderado do dossel por índice de vegetação (CVMVI) para estimativa de biomassa de soja a partir de imagens RGB baseadas em sistemas aéreos não tripulados. Revista ISPRS de fotogrametria e sensoriamento remoto: publicação oficial da Sociedade Internacional de Fotogrametria e Sensoriamento Remoto (ISPRS) , v. 27–41, 2019.

MARTIN, T. N. *et al.* **Tamanho ótimo de parcela e número de repetições em soja (*Glycine max* (L.) Merrill).** Ciência Rural, v. 35, n. 2, p. 271-276, 2005. doi: 10.1590/S0103-84782005000200004

MARTINELLI, L. *et al.* Soy expansion and socioeconomic development in municipalities of Brazil. **Land**, v. 6, n. 3, p. 62, 2017.

MARTINS, I. A. *et al.* **Lactofen and kinetin in soybean yield.** Pesquisa Agropecuária Tropical. v. 50, e64906, 2020. doi: 10.1590/1983-40632020v5064906

MATSUO, É. et al. Inheritance and genetic mapping of resistance to Asian soybean rust in cultivar TMG 803. **Crop Breeding and Applied Biotechnology**, v. 14, n. 4, p. 209–215, 2014.

MEOTTI, G. V. et al. Épocas de semeadura e desempenho agronômico de cultivares de soja. **Pesquisa Agropecuária Brasileira**, v. 47, n. 1, p. 14–21, 2012.

MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO – MAPA. **Estatística e dados básicos de economia agrícola agosto de 2013.** Disponível em:

<<https://www.gov.br/agricultura/pt-br/agroestatisticas/estatisticas-e-dados-basicos-de-economia-agricola/estatisticas-e-dados-basicos-de-economia-agricola-agosto-2013.pdf>>. Acesso em: 26 Jul. 2022.

MINISTERIO DO TRABALHO E EMPREGO - MTE - **Brasileiras e brasileiros no campo. Informações úteis**. Brasília, 2010. Disponível em: <<http://www.mte.gov.br>>. Acesso em: 26 de jul de 2022.

MIRANDA, F. D. **Produção, conteúdo de proteína e óleo no grão da soja: herdabilidades, correlações e seleção de genótipos superiores**. Dez. de 2006.

MOJENA, R. **Hierarchical grouping methods and stopping rules: Na evaluation**. Department of Management Science, University of Rhode Island, Kingston, Rhode Island 02881, USA. v. 20, n. 4, p. 359-363. 1977. <https://doi.org/10.1093/comjnl/20.4.359>

MORO, F. S. *et al.* **Produtividade de grãos em soja e seus componentes sob diferentes densidades de plantio**. Revista Tecno-Lógica, v. 25, n. 2, p. 314-319, 2021. doi: 10.17058/tecnolog.v25i2.16216

NOGUEIRA, A. P. O. et al. Análise de trilha e correlações entre caracteres em soja cultivada

NOGUEIRA, A. P. O. et al. Análise de trilha e correlações entre caracteres em soja cultivada nov./dez. 2012.  
nov./dez. 2012.

OLIVOTO, T. *et al.* **Multicollinearity in Path Analysis: A Simple Method to Reduce Its Effects**. Agronomy Journal, v. 109, p. 131–142, 2017. doi:10.2134/agronj2016.04.0196  
OLIVOTO, T. **Pliman: Tools for Plant Image Analysis**. R package version 1.1.0, <<https://CRAN.R-project.org/package=pliman>>. 2021.

PÁDUA, G. P. DE et al. Influência do tamanho da semente na qualidade fisiológica e na produtividade da cultura da soja. **Revista brasileira de sementes [Brazilian seed journal]**, v. 32, n. 3, p. 9–16, 2010.

PIEPHO, H. P.; MOHRING, J. **Computing heritability and selection response from unbalanced plant breeding trials**. Genetics, v. 177, n. 3, p. 1881-1888, 2007. doi: 10.1534/genetics.107.074229

PIMENTEL GOMES, F. **Curso de estatística experimental**. 15.ed., Piracicaba: FEALQ. 451p. 2009.

PÚBLIO JÚNIOR, E. *et al.* **Estimativas de parâmetros genéticos em genótipos de feijão-frade**. Revista de Ciências Agrárias, v. 41, n. 3, p. 806-814, 2018. doi: 10.19084/RCA17232

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2018. Disponível em: <https://www.R-project.org/>. Acesso em: jun. 2023.

RAMALHO, M. A. *et al.* **Aplicações da genética quantitativa no melhoramento de plantas autógamas**. Lavras, MG: UFLA. 2012.

RAMALHO, M. A. P.; SANTOS, J. B.; ZIMMERMANN, M. J. O. *Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro*. Goiânia: UFG, 1993. 271 p.

RESENDE, M. D. V.; DUARTE, J. B. **Precisão e controle de qualidade em experimentos de avaliação de cultivares**. *Pesquisa Agropecuária Tropical*, v. 37, n. 3, p. 182-194, 2007.

RIBEIRO, F. O. *et al.* **Genotypic Selection for Soybean Earliness**. *Journal of Agricultural Science*, v. 12, n. 11, p. 169-178, 2020. doi:10.5539/jas.v12n11p169

RICHETTI, A. **Viabilidade econômica da cultura da soja para a safra 2021/2022, em Mato Grosso do Sul**. [s.l.] EMBRAPA, 2021. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1133296/1/COT-262-2021.pdf> Acesso em: 19 set. 2023.

RIZZATTI ÁVILA, M. *et al.* **COMPONENTES DO RENDIMENTO, TEORES DE ISOFLAVONAS, PROTEÍNAS, ÓLEO E QUALIDADE DE SIEMENTES DE SOJA1**. Disponível em: <https://www.scielo.br/j/rbs/a/NYgNGWdMvPDqJ5Zxhpx9q3m/?format=pdf&lang=pt>. Acesso em: 16 set. 2023.

RODRIGUES, H. C. de A. *et al.* **Correlações genotípicas, fenotípicas e ambientais entre caracteres de mamoneira**. *Ciência e Agrotecnologia*, v. 34, n. 6, p. 1390-1395, 2010. doi: 10.1590/S1413-70542010000600005

ROTH, L. *et al.* **High-throughput field phenotyping of soybean: Spotting an ideotype**. *Remote Sensing of Environment*. v. 269, p. 112797, 2022. doi: 10.1016/j.rse.2021.112797

SACON, D. *et al.* Integration between genetic and chemical control on the progress of Asian soybean rust and yield. *Summa phytopathologica*, v. 46, n. 3, p. 198–204, 2020.

SAYAMA, T. *et al.* **Confirmation of the pleiotropic control of leaflet shape and number of seeds per pod by the Ln gene in induced soybean mutants**. *Breeding Science*, v. 67, n. 4, p. 363-369, 2017. doi: 10.1270/jsbbs.16201

SEDIYAMA, T.; SILVA, F.; BOREM, A. **Soja: do plantio a colheita**. UFV: [s.n.].

SILVA E. V. V. *et al.* **Genotypic and phenotypic parameters associated with early maturity in soybean**. *Pesquisa Agropecuária Brasileira*, v. 57, e02545, 2022. doi: 10.1590/S1678-3921.pab2022.v57.02545.

SILVA, F. *et al.* **Melhoramento da Soja**. Viçosa-MG: UFV. 2017.

SILVA, F. *et al.* **Soja: do plantio à colheita. 2ª edição**. Editora Oficina de Textos. Viçosa - MG: UFV. 2022.

SILVA, L. A. *et al.* **Estimating soybean leaf defoliation using convolutional neural networks and synthetic images**. *Computers and Electronics in Agriculture*. v. 156, p. 360-368, 2019. doi: 10.1016/j.compag.2018.11.040

SOARES, I. O. *et al.* **Genetic and phenotypic parameters associated with soybean**

**progenies in a recurrent selection program.** *Crop Breeding and Applied Biotechnology*, v. 20, n. 4, e28092046, 2020. doi: 10.1590/1984-70332020v20n4a59

SOUSA, C. A. F. *et al.* **Nova abordagem para a fenotipagem de plantas: conceitos, ferramentas e perspectivas.** *Revista Brasileira de Geografia Física*, v. 8, p. 660-672, 2015. doi: 10.26848/rbgf.v8.0.p660-672

TATTARIS, M.; REYNOLDS, M. P.; CHAPMAN, S. C. **A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding.** *Frontiers in Plant Science*, v. 7, p. 1-9, 2016. doi:10.3389/fpls.2016.01131

USDA. **Produção de Soja.** Disponível em: <https://www.usda.gov/>. Acesso em: 21 de jul de 2022.

VOGEL, J. T. *et al.* **Soybean Yield Formation Physiology – A Foundation for Precision Breeding Based Improvement.** *Frontiers in Plant Science*, v. 12, p. 719706, 2021. doi: 10.3389/fpls.2021.719706

WALTER, A.; LIEBISCH, F.; HUND, A. **Plant phenotyping: from bean weighing to image analysis.** *Plant Methods*, v. 11, n. 14, p. 1-11, 2015. doi: 10.1186/s13007-015-0056-8

WEI, M. C. F.; MOLIN, J. P. **Soybean Yield Estimation and Its Components: A Linear Regression Approach.** *Agriculture*, v. 10, n. 348, p. 1-13, 2020. doi:10.3390/agriculture10080348

WICKHAM, H. *et al.* **Welcome to the Tidyverse.** *Journal of Open Source Software*, v. 4, n. 43, p. 1686. 2019. doi: 10.21105/joss.01686

WRIGHT, S. **Correlation and causation.** *Journal of Agricultural Research*, Washington, v. 20, p. 557-585, 1921.

## ANEXOS

Tabela 1A - Estimativa de parâmetros genéticos de progênies de soja em análise individual de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG.

LOCAIS	PARÂMETROS	PROD	PMG	NGV	NVP	AT
LAVRAS 21/22	$\mu$	5784,44	162,32	2,64	103,55	76,33
	$\sigma_g^2$	638569,80	200,93	0,02	387,88	109,23
	$h_a^2$	0,23	0,45	0,46	0,27	0,26
	$r_{ap}$	0,48	0,67	0,68	0,52	0,51
	CVe	34,33	5,37	5,77	31,34	20,27
	LRT	0,08*	2,60e <sup>-19***</sup>	1,00e <sup>-05***</sup>	0,02**	8,00e <sup>-05***</sup>
IJACI 21/22	$\mu$	5492,30	164,04	2,67	123,26	72,76
	$\sigma_g^2$	838565,90	203,88	0,01	122,05	124,13
	$h_a^2$	0,25	0,47	0,47	0,07	0,27
	$r_{ap}$	0,50	0,68	0,68	0,26	0,52
	CVe	37,42	4,92	5,31	38,01	22,20
	LRT	0,01**	1,29e <sup>-21***</sup>	2,41e <sup>-07***</sup>	0,58 <sup>ns</sup>	0,00014***
LAVRAS 22/23	$\mu$	2449,85	158,84	2,42	145,62	74,24
	$\sigma_g^2$	520235,80	255,79	0,02	440,98	105,32
	$h_a^2$	0,41	0,50	0,33	0,2	0,31
	$r_{ap}$	0,64	0,71	0,57	0,44	0,55
	CVe	29,34	5,09	9,60	31,94	20,86
	LRT	9,89e <sup>-18***</sup>	3,66e <sup>-56***</sup>	6,67e <sup>-09***</sup>	0,002***	1,18e <sup>-08***</sup>
IJACI 22/23	$\mu$	4145,44	164,50	2,48	124,68	74,36
	$\sigma_g^2$	428549,80	205,03	0,03	501,46	113,94
	$h_a^2$	0,40	0,49	0,34	0,22	0,32
	$r_{ap}$	0,63	0,70	0,59	0,47	0,56
	CVe	16,68	5,07	8,54	35,84	21,28
	LRT	2,02e <sup>-16***</sup>	2,40e <sup>-44***</sup>	7,13e <sup>-10***</sup>	0,002***	3,53e <sup>-09***</sup>

Média geral ( $\mu$ ), variância genética ( $\sigma_g^2$ ), herdabilidade da média no sentido amplo ( $h_a^2$ ), acurácia seletiva ( $r_{ap}$ ), coeficiente de variação experimental (CV<sub>e</sub>) e teste da razão de verossimilhança (LRT). Produtividade de grãos (kg ha<sup>-1</sup>) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta (NVP), área dos trifólios (cm<sup>2</sup>) (AT).

<sup>ns</sup> não significativo ao nível de 0,10 de probabilidade, considerando a aproximação de qui-quadrado.

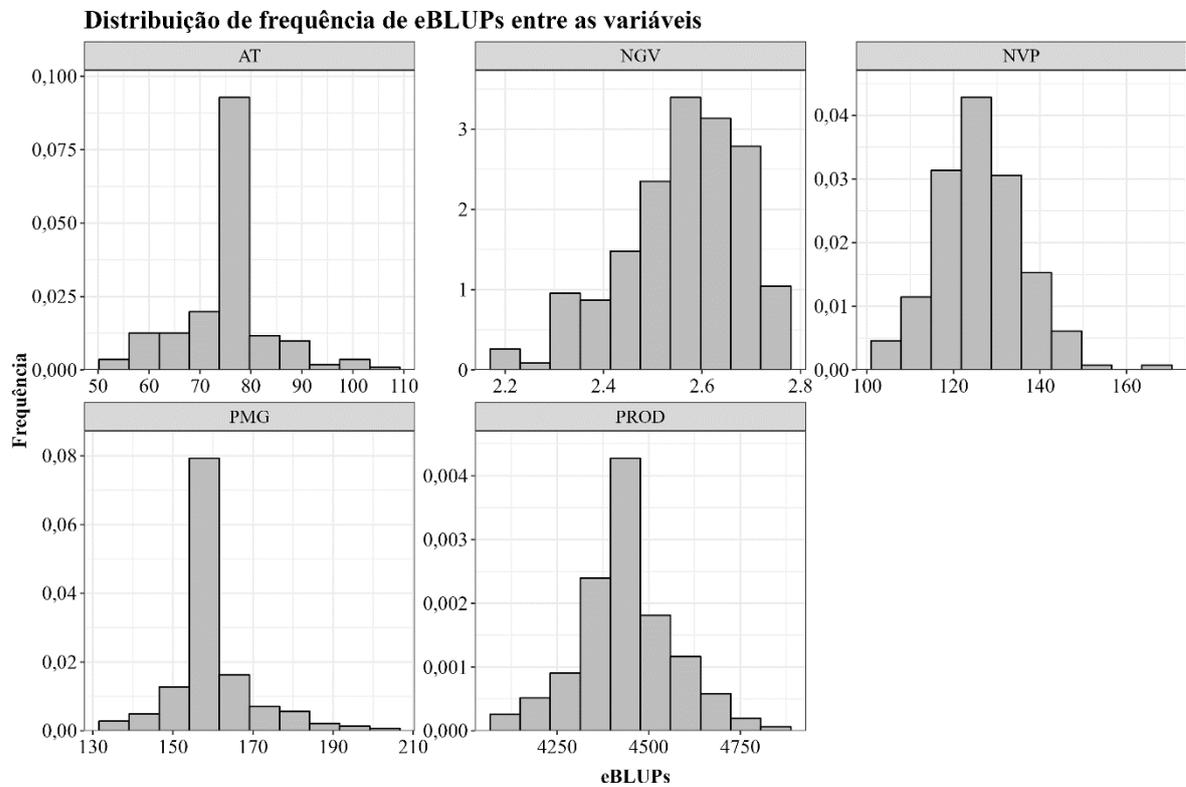
\*Significativo ao nível de 0,10 de probabilidade, considerando a aproximação de qui-quadrado.

\*\*Significativo ao nível de 0,05 de probabilidade, considerando a aproximação de qui-quadrado.

\*\*\*Significativo ao nível de 0,01 de probabilidade, considerando a aproximação de qui-quadrado.

Fonte: Do autor (2023).

Figura 1A – Distribuição de frequência de eBLUPs de progênies de soja em análise conjunta, de quatro ambientes, geração  $F_{2:3}$  na safra 21/22 e  $F_{2:4}$  na safra 22/23 em dois locais Lavras-MG e Ijaci-MG. Produtividade de grãos ( $\text{kg ha}^{-1}$ ) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta, área dos trifólios ( $\text{cm}^2$ ) (AT).

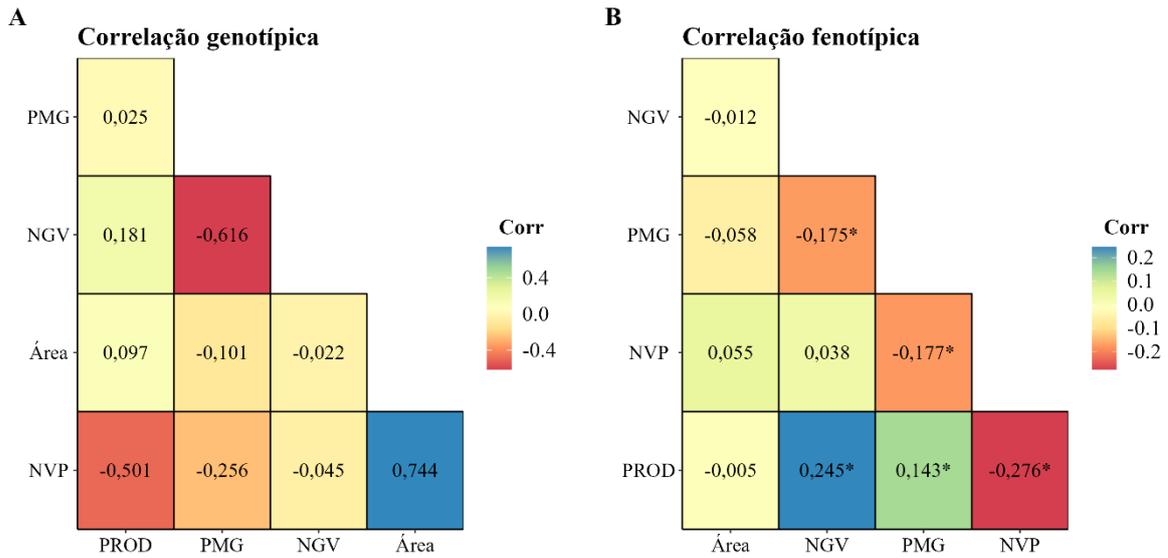


Fonte: Do autor (2023).

Figura 2A - Correlações genótípicas (A) e fenotípicas (B) de progênies de soja em análise conjunta, de quatro ambientes, geração  $F_{2:3}$  na safra 21/22 e  $F_{2:4}$  na safra 22/23 em dois locais Lavras-MG e Ijaci-MG. Produtividade de grãos ( $\text{kg ha}^{-1}$ ) (PROD), peso de mil grãos (gramas) (PMG), número de grãos por vagem (NGV), número de vagens por planta (NVP), área dos trifólios

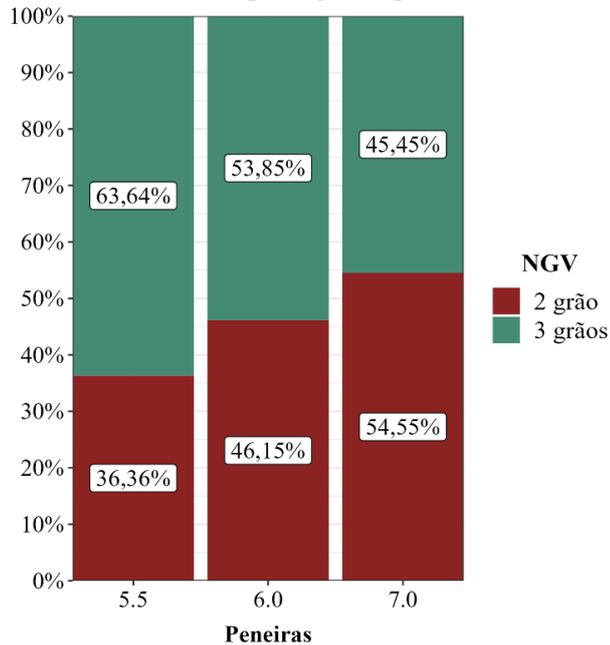
(cm<sup>2</sup>)

(AT).



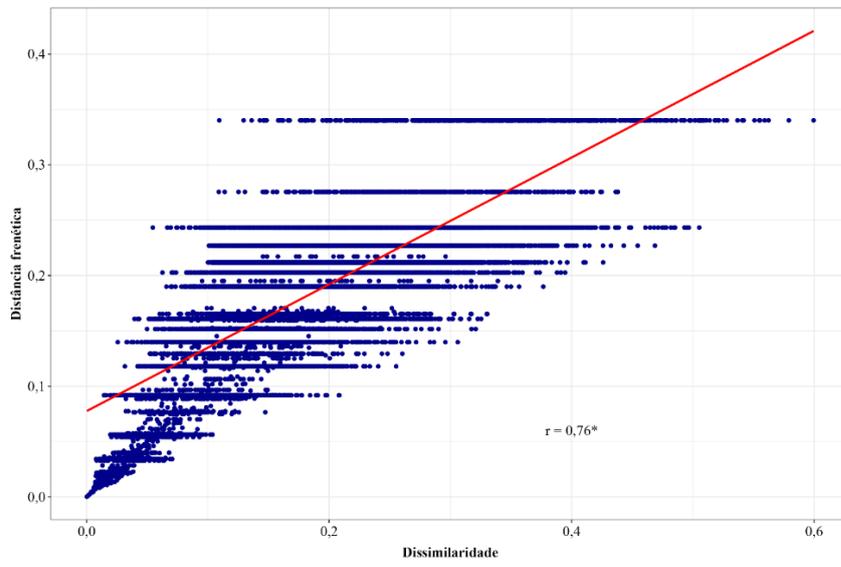
Fonte: Do autor (2023).

Figura 3A - Frequência de grãos por vagem e distribuição por peneira de progênes de soja em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG. Número de grãos por vagem (NGV).



Fonte: Do autor (2023).

Figura 4A. Distância frenética e dissimilariade de progênes de soja em análise conjunta, de quatro ambientes, geração F<sub>2:3</sub> na safra 21/22 e F<sub>2:4</sub> na safra 22/23 em dois locais Lavras-MG e Ijaci-MG.



Fonte: Do autor (2023).