



UNIVERSIDADE
ESTADUAL DE LONDRINA

GABRIELA INOCENTE

**MARCADORES AFLP E REDES NEURAIAS NA OBTENÇÃO
DE HÍBRIDOS INTERMEDIÁRIOS**

Londrina
2020

GABRIELA INOCENTE

**MARCADORES AFLP E REDES NEURAIAS NA OBTENÇÃO
DE HÍBRIDOS INTERMEDIÁRIOS**

Tese apresentada ao Programa de Pós-Graduação
em Agronomia da Universidade Estadual de
Londrina, como requisito parcial à obtenção do
título de Doutor em Agronomia.

Orientador: Prof. Dr. Paulo Maurício Ruas
Coorientador: Dr. Pedro Mario de Araújo

Londrina
2020

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

I35m Inocente, Gabriela.
Marcadores AFLP e redes neurais na obtenção de híbridos intermediários / Gabriela Inocente. - Londrina, 2020.
65 f. : il.

Orientador: Paulo Maurício Ruas.
Coorientador: Pedro Mario de Araújo.
Tese (Doutorado em Agronomia) - Universidade Estadual de Londrina, Centro de Ciências Agrárias, Programa de Pós-Graduação em Agronomia, 2020.
Inclui bibliografia.

1. Zea mays L. - Tese. 2. Híbridos Parcialmente Endogâmicos - Tese. 3. Marcadores Moleculares dominantes - Tese. 4. Redes Neurais Artificiais - Tese. I. Ruas, Paulo Maurício . II. Araújo, Pedro Mario de . III. Universidade Estadual de Londrina. Centro de Ciências Agrárias. Programa de Pós-Graduação em Agronomia. IV. Título.

CDU 63

GABRIELA INOCENTE

**MARCADORES AFLP E REDES NEURAIAS NA OBTENÇÃO
DE HÍBRIDOS INTERMEDIÁRIOS**

Tese apresentada ao Programa de Pós-Graduação em
Agronomia da Universidade Estadual de Londrina,
como requisito parcial à obtenção do título de Doutor
em Agronomia.

BANCA EXAMINADORA

Orientador: Prof. Dr. Paulo Maurício Ruas
Universidade Estadual de Londrina - UEL

Dra. Viviane Yumi Baba
Universidade Estadual de Londrina - UEL

Dr. Deoclécio Domingos Garbuglio
Instituto de Desenvolvimento Rural do Paraná -
IAPAR-EMATER

Dr. Nelson da Silva Fonseca Junior
Instituto de Desenvolvimento Rural do Paraná -
IAPAR-EMATER

Dra. Silvia Graciele Hulse de Souza
Universidade Paranaense- UNIPAR

Londrina, 3 de setembro de 2020.

AGRADECIMENTOS

À Universidade Estadual de Londrina e Programa de Pós-graduação em Agronomia da UEL, pela oportunidade de conhecimento e aprimoramento na área agrônômica. Agradeço também ao IAPAR por toda infraestrutura e suporte no desenvolvimento do projeto.

Agradeço ao meu orientador, Paulo Maurício Ruas, pela orientação neste trabalho, por sua afeição e dedicação pela ciência, e serenidade e retidão em tudo que faz.

Agradeço ao meu coorientador, Pedro Mário de Araújo, pelo suporte, amizade, direcionamento na área de pesquisa e acadêmica e por confiar no meu trabalho, desde a orientação do mestrado.

Ao Doutor Deoclécio Domingos Garbuglio, pelas orientações, incentivos e conhecimento durante toda a pós-graduação (Mestrado e Doutorado), mas acima de tudo, pelo exemplo na ciência e dedicação na área em que atua, por ser uma pessoa admirável, profissional exemplar e digna de toda amizade.

Ao meu marido, Rafael Delfini e a minha família, pela compreensão e pelo apoio. Por me incentivar na carreira acadêmica, científica e docente, por suportar as horas de ausências, as horas de mau humor e também por me ajudar em algumas atividades ao longo do doutorado.

Às amigas que a pós-graduação me deram, pela amizade, confiança, ajudas e desabafos: Bruna Góes, Amanda Letícia Pit Nunes, Maria Angélica Marçola, Alessandra Gonçalves e Renan Guilherme Vanzo.

Aos amigos e colegas de campo, desde o mestrado, ensinando, ajudando, incentivando e animando os dias longos de coletas de dados em campo, sem eles esse trabalho não seria possível: Antônio Alvez Ferreira, Luiz Gustavo Santos, Roberto Ribeiro dos Santos, João Bosco de Oliveira e Dirceu Duarte Araújo.

À Fundação Araucária pela concessão de bolsa de doutorado, sem a qual esse projeto não teria sido possível.

INOCENTE, Gabriela. **Marcadores AFLP e redes neurais na obtenção de híbridos intermediários**. 2020. 62 folhas. Tese (Doutorado em Agronomia) – Universidade Estadual de Londrina, Londrina, 2020.

RESUMO

Nos programas de melhoramento genético de milho para reconhecer genótipos com alto rendimento de grãos são necessários métodos de seleção eficiente que identifiquem as frequências alélicas e complementariedades das linhagens, de modo a maximizar a viabilidade do vigor híbrido. O uso de linhagens nas gerações iniciais de autofecundação (S_2) tendem a reduzir o número de cruzamentos teste, minimizando tempo e custos além de, apresentar alta produtividade de grãos e menor sensibilidade às variações ambientais. Buscando a otimização desse processo, o surgimento de marcadores moleculares e ferramentas computacionais garante a eficiência e resultados promissores. Os marcadores moleculares dominantes são ferramentas econômicas e altamente informativos para identificar genótipos associados as características de interesse, já para o modelo de aprendizado de máquinas *Multilayer Perceptron* (MLP) é possível a classificação e reconhecimento de padrões heteróticos de genótipos produtivos, mesmo quando a distribuição de probabilidade são desconhecidas. Além das tecnologias para identificações dos genótipos, o conhecimento meteorológico é fundamental para seleção do melhor híbrido intermediário. Nesse contexto, objetiva-se identificar padrões heteróticos a partir de Linhagens Parcialmente Endogâmicas (S_2) por meio de quatro combinações de primers AFLP e comparar com os seus respectivos cruzamentos (híbridos intermediários) em diferentes ambientes no Estado do Paraná. Os resultados obtidos pelos marcadores e os dados meteorológicos de cada ambiente testado, foram inseridos em um modelo de rede neural MLP, a fim de desenvolver um modelo de classificação voltado à seleção e descarte de genótipos de milho, visando alto rendimento de grãos e maximização do ganho genético. Os marcadores moleculares dominantes AFLP foram eficientes na identificação e clusterização dos padrões heteróticos mesmo em estágios de endogamia parcial (S_2) e a arquitetura de rede neural construída também foi eficiente na classificação de genótipos produtivos, visto que, a utilização conjunta de dados de marcadores moleculares e meteorológicos promoveram o aumento da capacidade preditiva do modelo.

Palavras chave: *Zea mays* L.; Híbridos Parcialmente Endogâmicos; Marcadores Moleculares dominantes; Redes Neurais Artificiais;.

INOCENTE, Gabriela. **AFLP markers and neural networks to obtain intermediate hybrids.** 2020. 62 pages. Thesis (PhD in Agronomy) – State University of Londrina, Londrina, 2020.

ABSTRACT

In maize breeding programs to recognize genotypes with high grain yield, efficient selection methods are needed to identify the allele frequencies and complementarities of the lines, in order to maximize the viability of the hybrid vigor. The use of lines in the initial generations of self-fertilization (S_2) tends to reduce the number of test crosses, minimizing time and costs, besides presenting high grain productivity and less sensitivity to environmental variations. Seeking to optimize this process, the rise of molecular markers and computational tools guarantees efficiency and promising results. The dominant molecular markers are economical and highly informative tools to identify genotypes associated with the characteristics of interest, for the Multilayer Perceptron (MLP) machine learning model, it is possible to classify and recognize heterotic patterns of productive genotypes, even when the distribution of probability are unknown. In addition to the technologies for genotype identification, meteorological knowledge is essential for selecting the best intermediate hybrid. In this context, the objective is to identify heterotic patterns from Partially Endogamic Lines (S_2) by means of four combinations of AFLP primers and compare with their respective crosses (intermediate hybrids) in different environments in the State of Paraná. The results obtained by the markers and the meteorological data of each tested environment, were inserted in a model of neural network MLP, in order to develop a classification model aimed at the selection and disposal of maize genotypes, aiming high grain yield and maximizing genetic gain. The dominant molecular markers AFLP were efficient in the identification and clustering of heterotic patterns even in stages of partial inbreeding (S_2) and the built neural network architecture was also efficient in the classification of productive genotypes since the joint use of molecular marker and meteorological data promoted the increase in the predictive capacity of the model

Key words: *Zea mays* L.; Partially Inbreeding Hybrids; Dominant Molecular Markers; Artificial Neural Networks.

LISTA DE FIGURAS

Figura 1 - Produção de milho por país.....6

Figura 2 - Representação de uma rede neural multi-layer perceptron.....19

ARTIGO A

Figura 1 - Porcentagem de ancestralidade dos parentais S_2 em relação aos híbridos intermediários de milho pelo software BAPS (Bayesian Analysis of Population Structure).....29

Figura 2 - Determinação do número de clusters (K) nos genótipos de milho pelo software BAPS (Bayesian Analysis of Population Structure).....30

Figura 3 - Neighbor-Net com base na distância de Nei & Li (1979), para todos os genótipos de milho (LPE0201, LPE0202 e HPE).....39

ARTIGO B

Figura 1 - Porcentagem de capacidade preditiva das classes genotípicas e padrões heteróticos em 100 repetições com reamostragem da rede neural.....42

Figura 2 - Análise das variáveis moleculares e meteorológicas **(a)** importância das variáveis em relação à classificação de produtividade. **(b)** Distribuição da importância das variáveis apresentada por cores azuis, sendo que, azul claro indicam maior importância e azul escuro menor, em relação à classificação rendimento de grãos.....43

Figura 3 - Dados de entrada: moleculares e agrometeorológico, marcadores moleculares e agrometeorológico **(a)** Distribuição das variáveis em barra, em relação a porcentagem de acerto através da MLP **(b)** BoxPlot das variáveis em relação a porcentagem de acerto pela rede MLP.....45

LISTA DE TABELAS

Tabela 1 - Tipo de híbridos e seus respectivos cruzamento.....	10
---	----

ARTIGO A

Tabela 1 - Descrição dos ambientes de avaliação dos 82 híbridos interpopulacionais, 19 linhagens S ₂ da população PC0201 e 23 linhagens S ₂ da população PC0202, no Estado do Paraná, Brasil. Safra 2016/2017.....	24
---	----

Tabela 2 - Quadrados médios obtidos por análise de variância conjunta, coeficiente de variação e médias obtidas dos Híbridos Experimentais de Linhagens Parcialmente Endogâmicas (HPE) para a variável produtividade de grãos (REND) em kg. ha ⁻¹ em três ambientes na safra 2016/2017.....	27
---	----

Tabela 3 - Agrupamento de médias por Scott-Knott a 5% de probabilidade dos vinte híbridos experimentais de linhagens parcialmente endogâmicas (HPE) mais produtivos em grãos (kg.ha ⁻¹), nas localidades de Londrina, Santa Tereza do Oeste e Guarapuava, PR na safra 2016/2017.....	27
---	----

Tabela 4 - Análise de Variância Molecular (AMOVA) das linhagens populações parcialmente endogâmicas (LPE) e híbridos experimentais de linhagens parcialmente endogâmicas (HPE).....	29
--	----

ARTIGO B

Tabela 1 - Descrição dos ambientes de avaliação dos 82 híbridos intermediários, 19 linhagens S ₂ da população PC0201 e 23 linhagens S ₂ da população PC0202, no Estado do Paraná. Safra 2016/2017.....	36
---	----

Tabela 2 - Classificadores dos períodos não críticos e críticos da cultura do milho baseados nos picos de temperaturas máximo e mínimo.....	37
--	----

Tabela 3 - Classificação de precipitação em milímetros (mm) das fases não críticas e críticas.....	38
---	----

Tabela 4 - Quadrados médios obtidos por análise de variância conjunta, coeficiente de variação e médias obtidas das Linhagens Parcialmente Endogâmicas (LPE) e Híbridos Experimentais de Linhagens Parcialmente Endogâmicas (HPE) para a variável produtividade de grãos (REND) em kg. ha ⁻¹ na safra 2016/2017.....	41
--	----

SUMÁRIO

1	INTRODUÇÃO	4
2	REVISÃO DE LITERATURA	6
2.1	ASPECTOS GERAIS DA CULTURA DO MILHO	6
2.2	MELHORAMENTO GENÉTICO EM MILHO	7
2.3	HIBRIDAÇÃO.....	9
2.4	HÍBRIDOS DE LINHAGENS PARCIALMENTE ENDOGÂMICAS	11
2.5	MARCADORES MOLECULARES NO MELHORAMENTO DE MILHO	13
2.6	MARCADORES AFLP	15
2.7	INTELIGÊNCIA ARTIFICIAL.....	17
2.8	MULTI-LAYER PERCEPTRON.....	18
3	ARTIGO A - AGRUPAMENTO E SELEÇÃO DE LINHAGENS PARCIALMENTE ENDOGÂMICAS DE MILHO COM O USO DE MARCADORES AFLP	21
4	ARTIGO B – CONSTRUÇÃO DE UMA REDE NEURAL MULTIPERCEPTRON APLICADA À CLASSIFICAÇÃO POR PRODUTIVIDADE DE GENÓTIPOS DE MILHO	33
5	CONSIDERAÇÕES FINAIS	47
	REFERÊNCIAS	48
	Referências – Artigo A.....	57
	Referências – Artigo B.....	60

1 INTRODUÇÃO

A cultura do milho (*Zea mays* L.) nos últimos anos tem registrado maiores incrementos na produção mundial (UNITED STATES DEPARTMENT OF AGRICULTURE - USDA, 2019). No Brasil a produtividade média de milho é de 97 milhões de toneladas distribuídas em 17 milhões de hectares (COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB, 2019). O aumento da produtividade deve-se ao grande avanço das pesquisas, por meio do melhoramento genético visando a adaptabilidade e estabilidade dos genótipos (MÔRO, 2018).

Em programas de desenvolvimento de híbridos de milho, o melhoramento de populações tem sido reconhecido como uma estratégia importante na obtenção de linhagens que resultem em combinações híbridas superiores. Isto se deve ao fato de que sucessivos ciclos de seleção aumentam gradativamente as frequências alélicas dos genes favoráveis na população (VENCOVSKY, 1987).

Os sucessivos ciclos de seleção são etapas que demandam um tempo maior na obtenção de linhagens totalmente endogâmicas (S_7 e S_8). Para acelerar esse processo uma das alternativas é a obtenção de híbridos intermediários obtidos por linhagens parcialmente endogâmicas na geração S_2 e S_3 , pois apresentam 75% e 87,5% respectivamente, dos locos em homozigose (CARVALHO; SOUZA; RAMALHO, 2010).

Diversos estudos apontam vantagens na obtenção de híbridos intermediários quando comparada com híbridos simples como: maior rapidez, menor sensibilidade às variações ambientais e menor interação com o ambiente (BARRETO et al. 2012; BERNINI; PATERNIANI, 2012; PRADO et al. 2013; KAGE et al. 2013; MIURA et al. 2014; RIGON, CAPUANI; RIGON, 2015; RAMIRÈZ; CARBALLO; AGUILAR, 2016; BEYENE et al. 2017).

O conhecimento prévio da diversidade do germoplasma e a boa capacidade geral de combinação (CGC) dos genitores, são fatores cruciais para o planejamento da produção de híbridos e conservação da variabilidade genética (SINGH et al. 2015; MENDOZA; NOLDIN, 2016), visto que, a identificação prévia das frequências alélicas auxilia na seleção dos cruzamentos em campo e aumenta a probabilidade de maximização do vigor híbrido.

Mediante o exposto, os marcadores moleculares dominantes como AFLP (Amplified Fragment Length Polymorphism) são ferramentas robustas que têm se mostrado eficientes e altamente informativos para estimar distância genética entre genótipos, a fim de obter a heterose (WU et al. 2019; MOSAFA et al. 2020; GIORDANI et al. 2019).

As principais vantagens da técnica de AFLP incluem o grande número de polimorfismos

por reação, isto é, explora variação ao longo de todo o genoma sem a necessidade de conhecimento prévio do genoma e baixo custo para execução (WU et al. 2019). Bespalhok, Guerra e Oliveira (2017) afirmam que as informações obtidas pelos marcadores moleculares combinado com avaliações dos genótipos em campo amplia a eficácia na identificação de padrões heteróticos de linhagens em fases iniciais (S₂).

Com o propósito de acelerar a eficiência das informações genéticas dos genótipos em teste nos programas de melhoramento genético, os modelos de aprendizado de máquinas computacionais têm ganhado notoriedade. As Redes Neurais Artificiais (RNAs) são constituídos por neurônios artificiais interconectados que apresentam retropropagação combinada e reconhecem padrões que contribuem na tomada de decisão em situações pré-determinadas (SHANMUGATHAN, 2016)

Atualmente, várias arquiteturas computacionais estão disponíveis nas RNAs, porém diversos estudos têm utilizado o modelo multilayer perceptron (MLP) na área de melhoramento genético em plantas e garantem eficiência do método (AZODI et al. 2019; JIANG et al. 2020; PARMLEY et al. 2019). Esse modelo é constituído por mais de um *perceptron* e a rede é supervisionada pelo algoritmo de retropropagação, ou seja, o próprio treinamento da rede ajusta os pesos e de bias para obtenção do mínimo erro possível (ABRISHAMI, SEPASKHAH, SHAHROKHANIA, 2019).

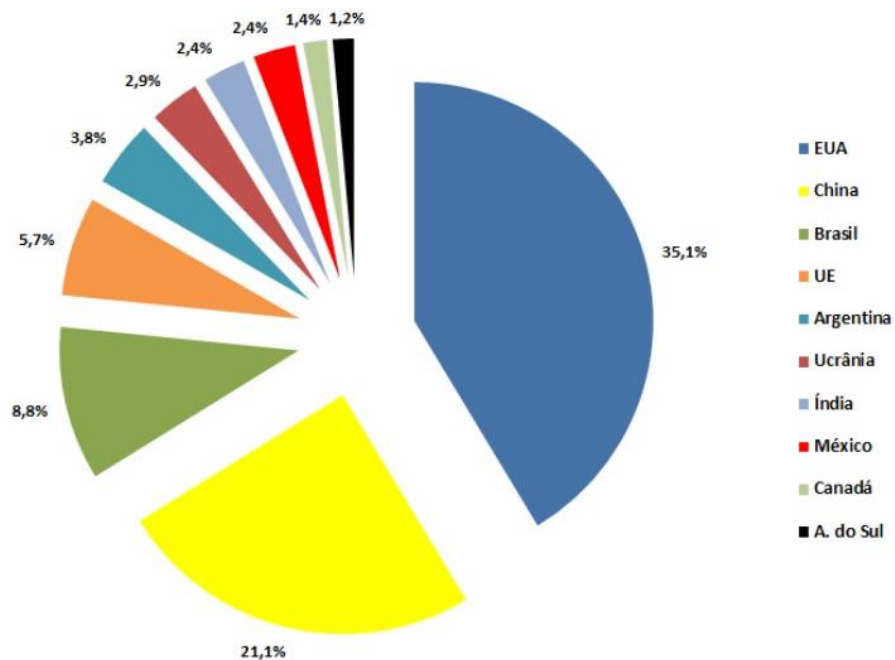
Na arquitetura do modelo MLP é possível incluir dados diversificados como características quantitativas e/ou qualitativas, marcadores moleculares, dados agrometeorológicos, sensores, imagens entre outros na camada de entrada. Já as informações obtidas na camada de saída podem ser classificadas no intuito de auxiliar na escolha de materiais, entre essas duas camadas encontram-se as camadas intermediárias ou camadas ocultas (FALK et al. 2020). Entretanto com a finalidade de se estimar genótipos de milho produtivos além dos dados obtidos em campo, as informações moleculares e meteorológicas aplicadas no modelo MLP são ideias para acurácia de seleção dos genótipos (LANA et al. 2017).

Assim, o presente estudo teve como objetivo identificar padrões heteróticos em grupos divergentes de linhagens parcialmente endogâmicas e suas aplicações na composição de híbridos intermediários de interesse agrônômico por meio de marcadores moleculares dominantes e os dados agrometeorológicos obter um padrão de classificação ou descarte de genótipos com foco em rendimento de grãos por meio de um modelo de redes neurais multicamadas treinadas .

2 REVISÃO DE LITERATURA

2.1 ASPECTOS GERAIS DA CULTURA DO MILHO

O milho (*Zea mays* L.) é o cereal de maior volume de produção do mundo, com aproximadamente 1 bilhão de toneladas obtidos na safra 2019/2020, conforme o relatório do USDA (United States Department of Agriculture). O Brasil é considerado o terceiro maior produtor de grãos de milho do mundo, ficando atrás apenas dos Estados Unidos e China, conforme a Figura 1 (USDA, 2019).



Fonte: FORMIGONI, I. 2018.

Figura 1. Produção de milho por país

Segundo o USDA (2019) o consumo mundial de milho deverá aumentar nos próximos anos, visto que, seu destino principal é a produção de ração para o uso animal. Porém o destino do milho abrange diversos segmentos como indústria, alimentícias, farmacêuticas e até na produção de etanol. No Brasil o Ministério da Agricultura, Pecuária e Abastecimento (MAPA) divulgou as projeções do agronegócio e o destino da produção de milho produzido na safra 2018/2019 a qual teve sua demanda voltada para consumo animal (51%), exportação (31%), consumo industrial (9%), consumo humano (2%), perdas (2%) e produção de semente (0,5%) e outros (4,5%).

Hoje em dia, a cultura de milho no Brasil representa aproximadamente 17 milhões de

hectares plantada com produtividade média de 98 milhões de toneladas distribuída em todo o território brasileiro, sendo cultivado em dois períodos denominados primeira e segunda safra. Conforme a Companhia Nacional de Abastecimento (CONAB), a safra 2018/2019 apresentou uma variação positiva de 3,8% em área planta e 17,5% em produtividade quando comparada com a safra 2017/2018.

No Brasil a produtividade de milho se concentra nos estados do Mato Grosso e Paraná com 32% e 17% respectivamente, e conforme as estimativas publicadas pela CONAB (2019) a cada safra esses estados apresentam uma variação positiva baixa em relação a área plantada (2,7% e 0,9% respectivamente) e uma variação positiva alta para produtividade (4,9% e 7,8% respectivamente). Esse aumento produtivo é necessário para suprir a demanda da cadeia produtora de proteína animal, principalmente, as oriundas de aves e suínos (DERAL, 2018).

Outro aspecto relevante quanto ao aumento da produtividade de milho se deve ao avanço no melhoramento genético das plantas, aumento da adubação, modernização dos maquinários agrícolas, utilização da irrigação e aprimoramento na agricultura de precisão (VIAN et al. 2016).

2.2 MELHORAMENTO GENÉTICO EM MILHO

O milho é considerado a espécie vegetal mais estudada, devido aos inúmeros métodos de melhoramento genético em plantas alógamas terem sido iniciados na cultura do milho (PATERNIANI; BERNINI, 2012). A espécie *Zea mays* L. é classificada como politípica, pois é cultivada em todos continentes devido ao grande número de genótipos adaptados (Fornasieri Filho, 2007).

Os primeiros relatos do melhoramento genético na cultura do milho, iniciou de forma inconsciente na seleção de características favoráveis, quando o homem deixou de ser nômade. Conforme Borém (1998) o melhoramento de plantas é “A arte e a ciência que visam as modificações gênicas das plantas para torná-las mais úteis ao homem”. Nesse sentido, o sucessivo plantio de plantas com características desejáveis provocou mudanças nas frequências gênicas, mas somente após a teoria de Mendel o aprimoramento do melhoramento genético progrediu.

O melhoramento genético desempenha um papel preponderante no desenvolvimento das plantas de importância para o homem, pois permite obter maior produtividade, uniformidade das plantas, adaptação a novos ambientes, qualidades nutricionais específicas, resistência a doenças e pragas, entre outros (MÔRO, 2018). Segundo Artuzo et al. (2019) o aumento da

produtividade na cultura do milho é resultado do nível tecnológico como plantio direto, manejo, tecnologia de precisão e as diversas características favoráveis selecionadas por meio do melhoramento de plantas.

Para obter cultivares de alta performance, programas de melhoramento genético de milho utilizam o melhoramento clássico para cruzamento dos genótipos testados, uma vez que, o milho é uma espécie alógama ou seja, não transfere seu genótipo para geração seguinte mas sim seus alelos. (HALLAUER; CARENA; MIRANDA FILHO, 2010). Assim, por meio de experimentos com repetições, classificação e seleção dos genótipos desejáveis os genótipos da população desejável são fixados (BUENO; MENDES; CARVALHO, 2006).

Atualmente, os programas de melhoramento de milho enfrentam consequências do uso frequente de cultivares de polinização aberta, motivo esse, que reduziu a variabilidade genética das populações base limitando o ganho genético. A fim de eliminar essas perdas os programas visam ampliar a base genética e reduzir a vulnerabilidade genética para obter ganhos de seleção. (ANDRADE, 2015). No contexto de se ampliar a base genética das fontes de extração de linhagens, estão envolvidos os processos de obtenção de novas populações que consistem em genótipos com locos homocigotos e heterocigotos, em que as estruturas genéticas dependem das frequências alélicas (MÔRO, 2018)

Para obter novas populações ou linhagens, Nass e Nishikawa (2001) afirmam que deve-se iniciar um processo denominado pré-melhoramento que consiste em:

- Síntese de novas populações base;
- Identificação de genes potencialmente úteis;
- Identificações de novos padrões heteróticos;
- Melhor conhecimento dos acessos *per se* e em cruzamentos;
- Maior quantidade de informações sobre o acessos;
- Auxílio no estabelecimento de coleções nucleares (*core collections*)

Para antecipar esses processos de pré-melhoramento dos programas, o uso da biotecnologia têm se tornado uma ferramenta imprescindível. A utilização de técnicas moleculares e modelos genético estatísticos ligados a genes de interesse, proporcionam uma seleção ágil, aumentando a probabilidade de obtenção de genótipos superiores (GUIMARÃES et al. 2018). Essa ferramenta molecular também possibilita estimar dentro de um grupo heterótico a divergência das linhagens em teste, promovendo a redução de cruzamentos em campo (CARENA; BARI, 2018).

De acordo com Sinimbu (2015) a identificação de grupos heteróticos na população base,

seja por marcadores moleculares ou outras ferramentas, proporciona a obtenção de híbridos de alta performance como maior produtividade, resistência a estresses bióticos e abióticos, adaptações a novos ambientes, menor tombamento e quebraimento de plantas, entre outros.

2.3 HIBRIDAÇÃO

As primeiras observações e descrições de hibridação em espécies vegetais foi relatada por Koelreuter (1766), Knight (1799) e Gartner (1849). Entretanto, Darwin em 1877, além de observar os efeitos de hibridação em plantas, conduziu experimentos comparando plantas autofecundadas com plantas cruzadas de milho e verificou o vigor híbrido, concluindo a superioridade das plantas obtidas da hibridação e perda de vigor em sucessivas autofecundações com base nos efeitos fenotípicos (GOULET; RODA; HOPKINS, 2017).

Após esses relatos, Beal em meados de 1880, cruzou variedades de polinização aberta de milho, observando a superioridade do híbrido em relação aos pais (KAPOOR; LATA; SHARMA, 2014), os mesmos resultados também foram obtidos por Sanborn (1890), McClure (1892), Morrow e Gardner (1893) e Willard, Weber e Clothier (1901).

A exploração dos híbridos para uso comercial teve início em 1909 por East e Shull, que relataram a perda de vigor das plantas quando essas passavam por sucessivas autofecundações quando cruzadas entre si e o desempenho do híbrido superava a média da população original. Shull conceituou heterose para os cruzamentos entre linhagens divergentes que apresentavam filhos com melhor desempenho do que a média de seus pais e endogamia para acasalamentos entre indivíduos aparentados (BERLAN, 2018). Entretanto, Shull, com esses novos conceitos, estipulou a necessidade da aquisição anual de sementes de milho, pois observou que ao cruzar linhagens aparentadas as sementes apresentavam perda de vigor devido a ocorrência da depressão endogâmica (HALLAUER; CARENA; MIRANDA FILHO, 2010).

A grande dificuldade de colocar em prática a hibridação era a manutenção e a baixa produtividade das linhagens, pois os custos inviabilizavam a produção de sementes. Assim, Jones em 1918, solucionou este problema com a produção de híbridos duplos, com boa aceitação agrônômica viabilizando a comercialização (SOUZA, 2018).

O primeiro milho híbrido comercializado, foi obtido no início da década de 1930, por meio dos experimentos de East e Shull, nos Estados Unidos, que foram substituindo gradativamente as variedades de polinização aberta (BESPALHOK; GUERRA; OLIVEIRA, 2018). Já no Brasil, os primeiros estudos com híbridos tiveram início na década de 1930, desenvolvendo cultivares mais adaptados às condições climáticas do país e no ano de 1939 o

Instituto Agrônomo de Campinas (IAC) foi pioneiro na comercialização de híbridos (SINIMBU, 2015).

Segundo Guimarães et al. (2018) a heterose tem sido altamente explorada nos programas de melhoramento de milho, identificando os genitores contrastantes e genes responsáveis pelas características ou complementariedades. Os genitores são populações divergentes que servirão de base para o desenvolvimento de linhagens endogâmicas, que posteriormente serão empregadas em cruzamentos (SILVA; MIRANDA, 2003).

As características dos híbridos que podem ser sintetizados foram relacionadas por Souza (2018) e são descritos a seguir:

Top-cross: as linhagens deste cruzamento podem ser de base genética ampla ou estreita. Esse tipo de híbrido é utilizado nos programas de avaliação de linhagens para utilização em híbrido;

Híbrido simples: Caracteriza-se por sua uniformidade e maior potencial produtivo, a semente tem um custo de produção mais elevado, devido à baixa produtividade da linhagem endogâmica utilizada como fêmea;

Híbrido simples modificado: Esse procedimento diminui o custo de produção de sementes, pois o genitor feminino apresenta um certo vigor que é manifestado em maior produção;

Híbrido Triplo: A linhagem polinizadora deve ser vigorosa, com grande quantidade de pólen, visto que, será intercalada ao híbrido, para garantir grãos satisfatórios nas linhas femininas;

Híbrido Duplo: Envolve quatro linhagens endogâmicas, apresentando menor vulnerabilidade e baixo custo da semente;

Híbrido intervarietal: Utiliza a heterose sem a necessidade da obtenção de linhagens, apresentam uma maior capacidade de adaptação devido à maior variabilidade genética em relação aos híbridos de linhagens. As gerações avançadas podem ser utilizadas como população base para o melhoramento interpopulacional.

A Tabela 1 apresenta os tipos de obtenção de híbridos e seus respectivos cruzamentos (BUENO; MENDES; CARVALHO, 2006):

Tabela 1. Tipo de híbridos e seus respectivos cruzamento

Tipos de híbridos	Cruzamentos envolvidos
<i>Top-cross</i>	Variedade x Linhagem

(Continuação)

Tipos de híbridos	Cruzamentos envolvidos
Híbrido simples (HS)	Linhagem (A)x Linhagem (B)
Híbrido simples modificado	(AxA') x B ou (AxA') x (BxB')
Híbrido triplo	(HS) x Linhagem (C)
Híbrido duplo	(HS) x (HS)
Híbrido múltiplo	(HS) x (HS) x (HS) x (HS)
Híbrido intervarietal	Variedade A x Variedade B

Além dos híbridos detalhados da Tabela 1, há os híbridos intermediários obtidos de linhagens com endogamia parcial ($0 < F < 1$), que serão detalhados no próximo tópico.

2.4 HÍBRIDOS DE LINHAGENS PARCIALMENTE ENDOGÂMICAS

Para obter um híbrido completamente endogâmico, as populações bases são submetidas a sucessivas autofecundações até obter linhagens homozigóticas ou puras, geralmente ocorre na sétima e oitava geração de autofecundação (S_7 e S_8) com a frequência em homozigose de 99,22% e 99,61% respectivamente. Hoje em dia, os híbridos comerciais de milho mais utilizados são os híbridos simples, pois são resultantes de duas linhas puras e possui maior potencial produtivo, em contrapartida, apresentam maior custo das sementes e suscetibilidade em ambientes que apresentam condições de estresses bióticos ou abióticos (GAYOSSO-BARRAGÁN et al. 2020; SOUZA et al. 2013).

Diante do exposto, o uso de linhagens parcialmente endogâmicas surge como uma alternativa promissora, por não ser necessário a condução à homozigose completa (S_7 e S_8), diminuindo o valor da semente híbrida e apresentar maior resistência quanto às oscilações climáticas quando comparada com híbrido simples (RAMIRÉZ; CARBALLO; AGUIAR, 2016). Além da diminuição do valor das sementes, as linhagens parcialmente endogâmicas diminuem os problemas relacionados à depressão por endogamia, vigor e queda de produtividade, que ocorre devido às sucessivas gerações de autofecundação (SOUZA JR., 2001).

Os pioneiros no estudo de híbridos intermediários Carlone e Russel (1998) apontam que linhagens S_2 apresentam maior produtividade do que as obtidas por linhagens completamente endogâmicas, além de ser realizado com maior rapidez, obter maior resistência às variações

ambientais adversas e menor interação genótipo por ambiente. O uso das progênie endogâmicas (S_1 ou S_2) é uma alternativa no melhoramento de populações para caracteres de baixa herdabilidade, uma vez que a endogamia aumenta a variância genética entre progênie e conduz ao aumento esperado por seleção (PARTENIANI; MIRANDA FILHO, 1987).

Conforme, Araújo e Paterniani (1999) a progênie S_2 apresenta $\frac{3}{4}$ da variância aditiva entre progênie e $\frac{1}{4}$ dentro da progênie. Sendo bom indicador para populações que já estão em seleção e nas populações que já apresentam redução na variância genética aditiva para os caracteres selecionados.

Para obter um bom híbrido intermediário, proveniente de linhagens parcialmente endogâmicas, os genitores devem apresentar boa capacidade geral de combinação (CGC), além disso, o conhecimento da diversidade do germoplasma é crucial para o planejamento da produção de híbridos e conservação da variabilidade genética. A CGC é o resultado de um dado genótipo produzir progênie com comportamento vantajoso quando cruzado com um conjunto de genitores divergentes, a caracterização das linhagens envolvidas no cruzamento é relevante pois está associada a genes de efeito aditivo (MEIRELLES, 2009; MURTADHA; ARIYO; ALGHAMDI, 2018). Conforme Bueno et al. (2006) a CGC de uma linhagem varia conforme a constituição genética do testador, o que pode resultar em cruzamentos com efeitos de dominância.

Outra prática de combinação complementar a CGC, é a capacidade específica de combinação (CEC) que se refere a uma combinação específica, podendo mudar o comportamento esperado com base na CGC. A CEC se associa a genes de efeito não aditivo conhecido como dominância e epistasia (PATERNIANI et al. 2008).

Diversos estudos com gerações S_2 de milho, onde os genótipos apresentam frequência homocigota de 75%, apontaram boa performance em cruzamentos e bom desempenho *per se* (CARVALHO, 2004; BARRETO et al. 2012; PRADO et al. 2013; GONÇALVES; DAVIDE 2013; MIURA et al. 2014; GARBUGLIO; ARAÚJO 2015; JUNIOR et al. 2017; PALAPARTHI; WALI, 2019).

Os resultados obtidos por meio de 90 progênie S_2 obtidos de duas populações de milho (BR 106 e IAPAR 26) em três locais (Londrina-PR; Ponta Grossa-PR; Rio Verde-GO) foi altamente eficiente para a obtenção de híbridos de alta produtividade, superando a média das testemunhas e dos parentais, representando um ganho de 4 a 5 anos no desenvolvimento quando comparado à obtenção de híbridos duplos ou triplos originados de linhagens totalmente endogâmicas (GARBUGLIO; ARAÚJO, 2004).

De acordo com Carvalho, Souza e Ribeiro (2003), foi observado que híbridos

intermediários, obtidos por meio de linhagens S_2 , apresentaram um potencial produtivo de 17% maior que as testemunhas e para as variáveis altura de plantas e espigas foram semelhantes, inferindo que os híbridos são vantajosos e têm capacidade de substituir híbridos comerciais.

Também foram realizados estudos com linhagens parcialmente endogâmicas (S_2) em regiões no México, a fim de obter híbridos intermediários em curto prazo, baixo valor da semente e adaptado ao clima chuvoso da região. Foram obtidas 174 sementes híbridas que posteriormente foram tratadas com campo magnético e tiveram acréscimo de 10 a 47% de germinação quando comparadas com sementes não tratadas, incentivando o uso de híbridos intermediários de milho e visando o aumento da renda dos produtores com baixa tecnologia mesmo situados em ambientes adversos ao desenvolvimento da cultura de milho, como é o caso da região em teste (RAMIRÈZ; CARBALLO; AGUILAR, 2016).

A fim de avaliar o potencial de linhagens S_2 , Junior et al. (2017) utilizaram uma linhagem-elite como testador nos cruzamentos *topcrosses* resultando em 108 híbridos intermediários em dois locais no Paraná (Guarapuava e Cândói) e dois em Minas Gerais (Paracatu e Guarda-Mor). Os híbridos intermediários apresentaram boa produtividade de grãos e estabilidade nos quatro ambientes, quando comparados com treze híbridos comerciais.

Conforme, Silva et al. (2017) foram observados que híbridos intermediários resultantes de linhagens parcialmente endogâmicas S_2 apresentaram alto potencial genético, desempenho e estabilidade dos materiais, concluindo que linhagens parcialmente endogâmicas são fontes promissoras e econômicas para obtenção de híbridos.

Por este motivo, o uso de linhagens S_2 além de apresentar uma boa combinação nos cruzamentos de milho resulta em híbridos com menor sensibilidade às variações ambientais, sendo uma fonte promissora para programas de melhoramento e rentável para os produtores.

2.5 MARCADORES MOLECULARES NO MELHORAMENTO DE MILHO

Os marcadores moleculares são ferramentas de grande potencial no melhoramento de plantas que auxilia na identificação de genótipos superiores em uma população segregante (BUSANELLO, 2016). Conforme Grattapaglia (2001) a introdução de técnicas da genética molecular permite a identificação, caracterização e mapeamento genético em um menor intervalo de tempo com maior segurança e eficiência, além de estimar a divergência genética entre os genitores e assim prever com precisão o desempenho da progênie.

Em meados de 1980, o primeiro marcador baseado em hibridação foi o marcador de Polimorfismo de Comprimento de Fragmentos de Restrição (RFLP). Esse marcador é

codominante e baseia-se na utilização de uma ou mais enzimas de restrição para o corte específicos da sequência de DNA, porém é uma técnica demorada, oneroso, exige uma sonda específica para hibridação e alta quantidade e qualidade do DNA (CAIXETA; FERRÃO; ZAMBOLIM, 2013).

Após o desenvolvimento da técnica de Polymerase Chain Reaction (PCR), foi possível obter a síntese enzimática de milhões de cópias de um segmento específico de DNA, revolucionando as técnicas de biologia molecular. O uso da PCR em análises genéticas permite uma melhor compreensão de processos de especiação, padrões de biogeografia em nível da espécie, bem como a estrutura genética de populações (FERREIRA; GRATTAPAGLIA, 1998).

O desenvolvimento da ciência genômica e do melhoramento genético molecular se deve ao desenvolvimento de marcadores moleculares cobrindo todo o genoma, mapas genéticos, automação no sequenciamento de DNA e o aperfeiçoamento da bioinformática (NADEEN et al. 2018)

Segundo, Borém e Caixeta (2009) a seleção assistida por marcadores moleculares no melhoramento genético se tornou uma ferramenta eficiente na transferência de alelos favoráveis. Para as características desejáveis em cada programa de melhoramento têm sido desenvolvidos marcadores moleculares ligados a essas características, permitindo a seleção indireta em gerações segregantes precoces (OSEI et al. 2018; NADEEN et al. 2018)

Conforme os estudos de Ferreira e Grattapaglia (1998), Frederizzi (1998), Ahmad et al. (2017) e Sharma e Sharma (2018) os marcadores moleculares apresentam inúmeras aplicações nos programas de melhoramento, destacando:

- Identificações de parentais;
- Proteção dos direitos do melhorista;
- Alocação de linhagens em grupos heteróticos;
- Teste de pureza de linhagens e híbridos;
- Mapeamento de caracteres qualitativos e quantitativos;
- Avaliação de Germoplasma (diversidade, classificação, filogenia);
- Introgessão de genes específicos;
- Pirimidização de genes;
- Seleção em geração precoce;
- Recuperação do parental recorrente;
- Manejo de genes

Segundo, Borém e Miranda (2013) os marcadores moleculares podem ser usados na

cultura do milho para avaliar a variabilidade genética, obter a predição de heterose com objetivo de selecionar genitores e híbridos adequados e diminuir o tempo de seleção em campo. O uso dos marcadores moleculares nos programas de melhoramento clássico de milho, além de reduzir o tempo na escolha de genitores proporciona melhoria no planejamento dos futuros cruzamentos (VETTORAZI, 2016).

No intuito de predição de híbridos com uso de marcadores moleculares, devem ser considerado: forte efeito de dominância; correlação negativa das frequências alélicas nos locos individuais das linhagens parentais; caracteres de alta herdabilidade; estreito intervalo de variação das médias de frequências dos alelos parentais; pelo menos 30 a 50% dos QTLs devem estar ligados aos marcadores moleculares (BERNARDO, 1992).

Nos estudos genéticos de plantas, há vários marcadores de DNA: RFLP (Restriction Fragment Length Polymorphism), RAPD (Random Amplified Polymorphism) AFLP (Amplified Fragment Length Polymorphism), SSRs (Simple Sequence Repeats) e, mais recentemente, os SNPs (Single Nucleotide Polymorfism ou Polimorfismo de Base Única). Para a escolha do marcador molecular adequado deve levar em consideração o desenvolvimento da técnica utilizada, avaliação e interpretação eficaz dos dados e análise dos dados (AMOM; NONGDAM, 2017).

Para estimar divergência entre linhagens de milho o uso de marcadores moleculares tem sido uma estratégia para minimizar as dificuldades de avaliação em campo, sendo uma delas o grande número de cruzamentos possíveis para a predição do desempenho de híbridos a fim de obter grupos heteróticos elite e híbridos produtivos (PATERNIANI et al. 2008).

A distância genética entre linhagens também pode ser estimada por marcadores moleculares. Falconer (1981) descreveu que a correlação entre divergência genética e heterose é positiva, pois conforme a teoria da genética quantitativa em qualquer grau de dominância superior (>0) a heterose é uma função da frequência alélica entre os genitores.

Nesse contexto, a distância genética entre linhagens de milho vêm sendo muito utilizada em programas de melhoramento, alocando grupos heteróticos (diferença na frequência de alelos das populações) e realizando a predição de cruzamentos superiores (XIA et al. 2005; LEGESSE et al. 2008; GRASSI et al. 2015; BEYENE et al. 2017, SHARMA; KUMAR, 2018).

2.6 MARCADORES AFLP

A técnica AFLP, segundo Vos et al. (1995), tem como base a amplificação seletiva por PCR de fragmentos de restrição derivados de uma digestão total do DNA genômico,

promovendo poderosa técnica de DNA *fingerprinting* para DNA de qualquer origem ou complexidade. Essa técnica envolve quatro etapas:

- 1- Clivagem do DNA genômico com duas enzimas de restrição, uma de corte raro e outra de corte frequente;
- 2- Ligação de adaptadores, que servem como sítios para o anelamento de *primers* e amplificação via PCR;
- 3- Etapas de PCR que consistem em uma seleção dos fragmentos amplificados, e divide em pré-seletivo (com *primers* contendo apenas um nucleotídeo arbitrário adicional na extremidade 3' do adaptador), e seletivo (com *primers* contendo outros dois nucleotídeos arbitrários adicionais, totalizando três);
- 4- Análise do subconjunto de fragmentos amplificados por eletroforese em gel de poliacrilamida ou em equipamentos automatizados por eletroforese capilar (FERREIRA; GRATTAPAGLIA 1998; CAIXETA; FERRÃO e ZAMBOLIM, 2013).

Dentre as tecnologias de marcadores moleculares, o marcador AFLP se destaca dos demais pois apresenta inúmeros marcadores analisados simultaneamente em uma única corrida eletroforética, ou seja, alto índice multiplex o que fornece grande número de fragmentos observados nas avaliações. Porém, apesar da vantagem de detecção de alta taxa de polimorfismo, o aparecimento de equívoco na interpretação das análises é comum, portanto, é necessário que sejam considerados os picos e bandas que sejam claros e de fácil visualização. Assim, fragmentos de difícil avaliação ou conflitantes devem ser preteridos (CAIXETA; FERRÃO e ZAMBOLIM, 2013).

O AFLP é considerado uma técnica eficiente para estimar a diversidade e distância genética em milho associados com predição de híbridos, conforme estudos comparativos (QI et al. 2010; MOLIN et al., 2013; ROY; KIM, 2016; GIORDANI et al. 2019).

Essa técnica também constitui estudos relacionados à condições ambientais e diversidade genética (MAKUMBI et al. 2011), detectar introgressões genéticas (TSEHAYE et al. 2016; RITTE et al. 2018) e construir mapas de ligações genético (VUYLSTEKE et al. 1999; SU et al. 2017; GIORDANI et al., 2019).

As principais vantagens dos marcadores moleculares AFLP incluem o baixo custo para execução e o grande número de polimorfismos por reação, sem a necessidade de conhecimento prévio do genoma. Também é característico da técnica de AFLP, gerar marcadores espécie-específicos, capazes de identificar espécies relacionadas distinguindo-as umas das outras e ainda, caracterizar geneticamente cultivares de uma mesma espécie (BRAMMER, 2000; ROY; KIM, 2017).

Qi et al. (2010) avaliaram a distância genética de 36 híbridos provenientes de um cruzamento dialélico com nove linhas puras de milho. A divergência genética foi estimada por marcadores moleculares AFLP e comparada com heterose em campo, apresentando uma correlação moderada (0,4096). O marcador AFLP tem alto polimorfismo e pode ser usado para detectar divergência genética e identificar cruzamentos entre linhagens com valores positivos para Capacidade Geral de Combinação (CGC) e heterose.

A variabilidade genética na cultura do milho é imprescindível para obter materiais produtivos, desse modo, Molin et al. (2013) avaliou a distância genética de 48 variedades de milho no Paraná e Rio Grande do Sul, com três marcadores moleculares (AFLP, SSR e RADP). O resultado identificou alta variabilidade entre os materiais, independente do marcador utilizado, entretanto, o marcador AFLP foi mais eficiente por apresentar maior número de fragmentos polimórficos em relação aos outros marcadores testados.

Segundo, Giordani et al. (2019) ao analisarem 145 acessos de germoplasma de linhagens de milho testaram evidências preliminares de associação entre a doença ferrugem polissora (*Puccinia polysora*) que é causadora de perdas significativas na produção e marcadores AFLP. O estudo detectou a existência de três marcadores AFLP que está associado a resposta do milho em relação a ferrugem, corroborando com a alta eficiência da técnica AFLP.

2.7 INTELIGÊNCIA ARTIFICIAL

A população mundial deverá crescer cerca de 33% nas próximas década, atingindo quase 10 bilhões de habitantes em 2050, de acordo com dados da FAO (2019), esse incremento populacional deverá impactar principalmente nos recursos finitos, entre eles o aumento da demanda de alimentos.

Com esse impacto, até 2050 serão necessários aumentar a produção dos alimentos em 70% da produção atual, visto que, o uso de recursos para produção como energia, fertilizante e pesticida deverão ser reduzido, caminhando para uma agricultura sustentável (FAO, 2019). Essas modificações no setor agrícola vêm empregando novas tecnologias, entre elas, o uso computacional uma ferramenta promissora e benéfica para a agricultura.

Com o avanço das transformações agrícolas, o conhecimento das mudanças climáticas é de extrema importância, principalmente no desenvolvimento de novas variedades com maior estabilidade de produção. (CECCARELLI et al. 2010; LANA et al. 2017). Porém, para desenvolver genótipos ideais em condições climáticas divergentes, têm se notado uma maior frequência no uso de Inteligência Artificial (IA).

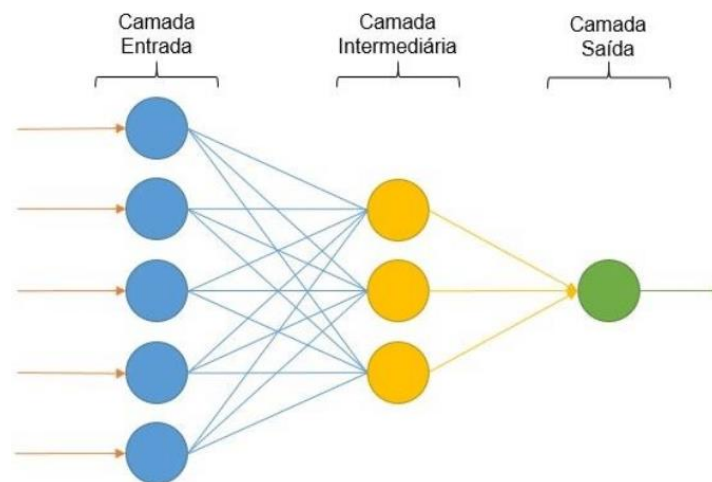
Conforme Jha et al. (2019) a IA é uma ferramenta computacional que indica a melhor decisão a ser tomada em uma determinada situação por meio de algoritmos. Esse método consiste em treinar uma rede computacional, com a inclusão de dados, e torná-lo automatizado. Assim, a rede com experiências acumuladas, forma uma habilidade cognitiva semelhante ao do ser humano.

Entende-se que, a inteligência artificial e seu ramo de *Machine learning* (ML), têm proporcionado avanços significativos em diversos ramos de atuação de forma eficiente. Os modelos de Redes Neurais Artificiais (RNAs), derivados da ML, estão sendo aplicada de forma eficaz no contexto do melhoramento genético de milho, pois reconhecem padrões e apresentam retropropagação combinada através de neurônios artificiais interconectados (GARCÍA-MARTINÉZ et al. 2020).

Segundo An et al. (2019), ao avaliar genótipos de milho tolerantes a estresse por seca através de redes convolucionais, observaram 98,14% e 96% na acurácia do modelo para identificação e classificação dos genótipos respectivamente. Já Qin et al. (2018) buscaram explorar a predição da taxa ótima de nitrogênio para milho com a aplicação das metodologias de ML, foram utilizados dados de 47 experimentos da região americana produtora de milho (Corn Belt) e observaram que a melhoria nas estimativas de época, solo e status hidrológico parecem ser essenciais para o sucesso na modelagem da demanda de nitrogênio pela cultura do milho.

2.8 MULTILAYER PERCEPTRON

A rede neural multilayer perceptron (MLP) é uma das arquiteturas mais importante disponíveis dentro das redes neurais artificiais (RNAs), pois é um modelo de aprendizado profundo envolvendo mais de uma camada capaz de resolver problemas complexos. Conforme Batista (2012) esse modelo é composto por camadas de neurônios denominadas: entrada, intermediárias ou ocultas e saída conforme a Figura 2.



Fonte: ABEL, F. 2016

Figura 2. Representação de uma rede neural multilayer perceptron

As arquiteturas das MLP são classificadas como supervisionada, ou seja, o modelo aprende os valores ideais de pesos e bias pelo algoritmo de retropropagação (*error back-propagation*) com objetivo de ajuste de parâmetro e reduzir erros. Após o aprendizado é possível inserir novos dados na camada de entrada a fim de observar a resposta expressa na camada de saída (OLIVEIRA, 2019).

Para o treinamento da rede MLP é utilizado o algoritmo chave *backpropagation* que é denominado de “diferenciação no modo reverso” (STORM et al. 2019), isto é, o modelo é capaz de ajustar os parâmetros à medida que o fluxo de sinal se move entre as camadas em uma constante ida e volta, até que o erro não possa mais ser reduzido, esse processo é conhecido como convergência (UECE, 2017).

As MLP estão presentes em diversas áreas para resolução de problemas das mais variadas vertentes, como a área médica (WANG et al. 2018), econômica (ALAMEER et al. 2019), industrial (SJARIF et al. 2020), biológica (ZYKOVA et al. 2020), agricultura (PARMLEY et al. 2019), entre outras.

Atualmente, no campo de melhoramento genético os trabalhos envolvendo MLP ainda são incipientes, porém as pesquisa mostram eficiência do modelo em estimar valores genéticos (SILVA et al. 2016), avaliar diversidade genética (JIANG et al. 2020), identificar genótipos superiores (PARMLEY et al. 2019) e seleção genômica (GONZÁLEZ-CAMACHO et al. 2018). Os métodos de MLP no melhoramento otimiza e garante avanços quanto aos ganhos genéticos em novas linhagens parentais, híbridos ou variedades comerciais para diferentes características de interesse, principalmente produtividade.

Diversos estudos apontam que o ganho genético por meio de treinamento supervisionado tem maior acurácia quando é utilizado um maior conjunto de dados, sejam eles, características quantitativas e/ou qualitativas, marcadores moleculares e até dados meteorológicos (FALK et. 2020; PARMLEY et al. 2019). Em programas de melhoramento genético em milho a predição, classificação de genótipos de alta performance e identificação de grupos heteróticos são essenciais para o sucesso (MEENA et al. 2017).

Nesse sentido, o desenvolvimento de um modelo de MLP no intuito de classificação para seleção ou descarte de genótipos com alto rendimento de grãos é imprescindível para ganhos genéticos em um determinado *pool* genético.

3. ARTIGO A – AGRUPAMENTO E SELEÇÃO DE LINHAGENS PARCIALMENTE ENDOGÂMICAS DE MILHO COM O USO DE MARCADORES AFLP

Resumo

Os programas de melhoramento de milho atualmente são pautados pela busca de genótipos superiores no menor espaço de tempo possível. Nesse sentido, informações de linhagens relativas à capacidade de combinação ainda em estágios de endogamia parcial podem acelerar o processo de obtenção de híbridos superiores, bem como aumentar a probabilidade de maximização da heterose ou vigor híbrido quando em cruzamentos. Esse direcionamento dos processos de seleção ou descarte podem ser intermediados por marcadores moleculares no sentido de auxiliar a identificação dos grupos ou padrões heteróticos. Assim, o presente trabalho buscou acessar informações relativas aos padrões heteróticos em grupos divergentes de linhagens parcialmente endogâmicas e suas aplicações na composição de híbridos intermediários de interesse agrônomico com uso de marcadores AFLP. O resultado foi comparado com os dados de campo das linhagens avaliadas e seus cruzamentos *per se* (82 híbridos intermediários) em ambientes distintos no Estado do Paraná. Foi realizada uma análise conjunta seguido de agrupamento de médias pelo método de Scott-Knott a 5% de probabilidade. Foi observado um padrão de performance de determinados híbridos associados a parentais comuns que se repetiam ao longo dos ambientes, mesmo que inseridos em híbridos distintos, sendo os parentais machos (PC0201): 103, 107, 109, 111, 113, 118 e 119, e os parentais fêmeas (PC0202): 204, 215, 220 e 223. Por meio da análise bayesiana ficou evidente que a ancestralidade de LPE0202 foi superior em relação a LPE0201 nos híbridos intermediários. Observou-se através da neighbor-Net que cruzamentos de linhagens contrastante geneticamente resultaram em híbridos com alto rendimento de grãos, confirmando a relação positiva entre produtividade dos materiais avaliados e distância genética obtidas por marcadores moleculares AFLP. Os marcadores moleculares dominantes AFLP foram eficientes na identificação e clusterização dos padrões heteróticos mesmo em estágios de endogamia parcial (S₂). A conservação dos padrões heteróticos no estágio S₂ permite um descarte ou seleção eficiente de linhagens promissoras, para fases avançadas de endogamia. O uso de híbridos comerciais oriundos de um “*pool*” genético comum para a formação das populações originais, foram evidenciados pelo agrupamento de genótipos distintos em um mesmo cluster.

Palavras-chave: Marcador AFLP; endogamia incompleta; híbridos intermediários

Abstract

Currently, maize breeding programs are guided by the search for superior genotypes in the shortest possible time. In this sense, information from lines regarding the combining ability still in stages of partial inbreeding can accelerate the process of obtaining superior hybrids, as well as increase the probability of maximizing heterosis or hybrid vigor when in crosses. This direction of the selection or discards processes can be mediated by molecular markers in order to help the identification of heterotic groups or patterns. Thus, the present work sought to access information related to heterotic patterns in divergent groups of partially endogamic lines and their applications in the composition of intermediate hybrids of agronomic interest using AFLP markers. The result was compared with the field data of the lines evaluated and their crosses *per se* (82 intermediate hybrids) in different environments in the State of Paraná. A joint analysis was performed followed by grouping of means by the Scott-Knott method at 5% probability. It was observed a performance pattern of certain hybrids associated with common parents that were repeated throughout the environments, even if inserted in different hybrids, being the male parents (PC0201): 204, 215, 220 and 223, and the female parents (PC0202): 103, 107, 109, 111, 113, 118 and 119. Through Bayesian analysis it was evident that the ancestry of LPE0202 was superior in relation to LPE0201 in the intermediate hybrids. It was observed through the neighbor-Net that crosses of genetically contrasting lines resulted in hybrids with high grain yield, confirming the positive relationship between productivity of the evaluated materials and genetic distance obtained by molecular markers AFLP. The dominant molecular markers AFLP were efficient in the identification and clustering of heterotic patterns even in stages of partial endogamic (S_2). The conservation of heterotic patterns in the S_2 stage allows an efficient discard or selection of promising lines, for advanced stages of inbreeding. The use of commercial hybrids from a common genetic "pool" for the formation of the original populations, was evidenced by the grouping of distinct genotypes in the same cluster.

Keywords: AFLP Marker; incomplete inbreeding; intermediate hybrids

Introdução

Na cultura do milho, a identificação prévia das frequências alélicas auxilia na seleção dos cruzamentos em campo. Essa identificação e a complementariedade genética das linhagens tende a reduzir o número de cruzamentos testes, além de direcionar as seleções nas gerações iniciais de autofecundação, como S_1 e S_2 , de modo a aumentar a probabilidade de maximização do vigor híbrido ou heterose, manifestada nos cruzamentos.

Entretanto, a interação entre divergência genética dos parentais e desempenho do híbrido só é considerada positiva, quando efeitos de dominância (ou não aditivos) estão presentes e os parentais envolvidos apresentarem diferenças nas frequências alélicas (GADHERI; ADAMS; NASSIB, 1984), sendo que a expressão dos efeitos não aditivos vem sendo estudados nas linhagens em estágios superiores de endogamia ($S_7 - S_n$) e em combinação nos híbridos derivados destas.

No entanto, diversos estudos observaram vantagens quanto ao uso de linhagens parcialmente endogâmicas (CARLONE; RUSSEL, 1988; RIGON; CAPUANI; RIGON, 2015; RAMIRÈZ; CARBALLO; AGUILAR, 2016; BEYENE et al. 2017) em combinações, obtendo híbridos intermediários com menor sensibilidade às variações ambientais, menor interação com ambiente e agilidade na obtenção, quando comparado aos híbridos simples proveniente de linhagens totalmente endogâmicas.

O direcionamento de cruzamentos em geral é realizado considerando os padrões heteróticos que as linhagens a serem testadas pertencem, e são utilizados testadores divergentes em relação a esses padrões, podendo estes serem determinados através do uso de marcadores moleculares ou via cruzamentos top-cross entre as linhagens e os diferentes testadores.

Os marcadores moleculares tradicionais como AFLP (Amplified Fragment Length Polymorphism) têm se mostrado eficientes e econômicos, bem como altamente informativos em identificar genótipos associados com caracteres de interesse ou determinações de padrões heteróticos em genótipos com estágios avançados de endogamia (SHARMA et al. 2018; GIORDANI et al. 2019, WU et al. 2019; MOSAFA et al. 2020). No entanto a combinação entre marcadores moleculares e avaliações a campo, tem se mostrado altamente relevantes para identificações de padrões heteróticos em fases iniciais (S_2) a fim de explorar heterose na obtenção de híbridos intermediários (BESPALHOK; GUERRA; OLIVEIRA, 2017)

Dessa forma, o presente estudo teve o objetivo de acessar informações relativas aos padrões heteróticos em grupos divergentes de linhagens parcialmente endogâmicas por meio de marcadores AFLP e suas aplicações na composição de híbridos intermediários de interesse agrônômico.

Material e Métodos

Experimentos

Foram utilizadas duas populações de milho PC 0201 e PC 0202, sintetizadas no Instituto de Desenvolvimento Rural do Paraná -IAPAR-EMATER por meio do inter cruzamento entre híbridos comerciais, na safra 2005/2006, a qual apresentaram alta capacidade específica de

combinação (CEC) em diferentes variáveis e produtividade de grãos estatisticamente similar à testemunha utilizada (BRS 3003).

As 90 linhagens S₂ de cada população foram submetidas a um cruzamento dialélico parcial circulante interpopulacional (CDPCI), conforme metodologia de Miranda Filho e Vencovsky (1999). Por meio desse método de cruzamento cada progênie de PC0201 foi cruzada com 4 progênies de PC0202, resultando em 360 híbridos avaliados na safra 2012/2013 em dois locais (Londrina e Guarapuava, no estado do Paraná).

Posteriormente foram selecionadas as linhagens parcialmente endogâmicas (LPE) *per se* de cada população (19 LPE originada da PC0201 e 23 LPE proveniente da PC0202) e 82 híbridos intermediários parcialmente endogâmicos (HPE) com maior produtividade de grãos e menor índice de acamamento de plantas.

Locais e Avaliações

Os ensaios envolvendo 82 híbridos intermediários foram conduzidos nos municípios de Londrina, Guarapuava e Santa Tereza do Oeste, no estado do Paraná, juntamente com duas testemunhas comerciais: DKB210 e 2B810 (Tabela 1).

Tabela 1. Descrição dos ambientes de avaliação dos 82 híbridos intermediários parcialmente endogâmicos (LPE) no Estado do Paraná, Brasil. Safra 2016/2017.

Ambiente (Código)	Altitude	Latitude	Longitude	* Clima	Solo
Londrina (LD)	569 m	23°22'0.06"S	51° 9'55.61"O	** Cfa	Latossolo Vermelho distrófico
Santa Tereza do Oeste (STO)	749 m	25° 08'S	53° 58'O	*** Cfa	Latossolo Vermelho distroférrico típico
Guarapuava (GUA)	1038 m	25°23'51.87"S	51°32'35.25"O	*** Cfb	Latossolo Bruno Álico e associações com Cambissolo Álico

* Segundo o sistema de classificação climática de Köppen-Geiger. ** Cfa: clima subtropical úmido. *** Cfb: clima subtropical úmido mesotérmico

Os ensaios foram avaliados em blocos casualizados com duas repetições por local, sendo a parcela experimental constituída de uma linha de 5 metros de comprimento e 0,80 m entre linhas, deixando-se 5 plantas por metro linear após desbaste. A variável analisada foi rendimento de grãos (REND), obtida a partir da pesagem dos grãos de todas as espigas colhidas na área útil das parcelas (4m²), ajustado para kg. ha⁻¹ e corrigido para umidade padrão de 14,5%.

Marcador Molecular - AFLP

Para as análises moleculares foram extraídas o DNA de cada genótipo (42 linhagens S₂ e 82 HPE) após 10 dias de emergência em campo, mediante o procedimento proposto por Doyle e Doyle (1987) com modificações. Onde o CTAB foi substituído por MATAB (Mixed Ayltrimethylammonium Bromide, Sigma) no tampão de extração. As amostras de DNA foram quantificadas no Nanodrop (Thermo Fisher Scientific, USA) para a verificação da concentração e pureza das mesmas.

Cerca de 700 ng do DNA genômico foram digeridos utilizando simultaneamente enzimas de restrição de corte frequente MseI (1U) e corte raro EcoRI (5U) a 37 °C por 4h, em seguida, os fragmentos gerados foram ligados a adaptadores específicos pela T4 DNA ligase (2U) por 22°C a 1h e incubada a 70 °C por 1h, para inativação térmica das enzimas de restrição (Vos et al. 1995).

A amplificação pré-seletiva foi realizada utilizando 3,5 µL do kit GoTaq® Green Master Mix (Promega, Winchester, EUA), 4,75mM do primer pré-seletivo, 3 µL da diluição da reação de restrição ligação e 2,92 µL de água ultrapura para completar o volume de 10 µL. O programa de amplificação pré-seletiva em termociclador consiste em: 1 ciclo de 72°C por 2 min, 20 ciclos de 90°C por 1s, 56°C por 30s e 72°C por 2 min, seguido de 1 ciclo final de 60°C por 30 min. Em seguida, um volume de 5 µL da reação pré-seletiva foi diluído em 45 µL de água ultrapura.

Para a amplificação seletiva dos genótipos de milho, foram utilizadas quatro combinações de *primers* marcados com fluoróforos *EcoRI* ACG (FAM-azul) / *MseI*-CAG; *EcoRI* AGG (PET-vermelho) / *MseI*-CTG; *EcoRI* AAG (NED-amarelo) / *MseI* CTG e *EcoRI* AAG (HEX-verde) / *MseI*-CTC). A reação de amplificação seletiva foi composta dos *primers* seletivos *MseI* 5 µM e *EcoRI* 5 µM, *GoTaq® Green Master Mix* (Promega, Winchester, EUA) 3,5 µL e 2,5 µL da reação da pré-amplificação diluída e 2,92 µL de água ultrapura para completar o volume de 10 µL. O programa para amplificação em termociclador foi composto por 1 ciclo de 94°C por 2 min, 65°C por 30 s e 72°C por 2 min; 8 ciclos de 94°C por 1 s, 64°C por 30 s e 72°C por 2 min; 23 ciclos de 94°C por 1 s, 56°C por 30 s e 72°C por 2 min e 1 ciclo final de 60°C por 30 min.

Para resolução dos produtos de PCR obtidos da reação seletiva foi feito um sistema multiplex com a seguinte diluição: 2,0 µL de FAM, 2,0 µL de NED, 2,0 µL de VIC, 2,0 µL de PET com 3 µL de água ultrapura. A partir desta mistura, adicionou-se 1,0 µL da mesma a 0,2 µL de size standard LIZ 500 (Applied Biosystems, Califórnia, EUA) e 8,8 µL de formamida Hi-Di Pop-7 (Applied Biosystems, Califórnia, EUA), desnaturado a 95°C por 3 minutos e

imediatamente resfriado em gelo. A eletroforese capilar foi realizada com um sistema automatizado ABI 3500XL (Applied Biosystems, Califórnia, EUA).

Análise estatística

Foi realizado a análise de variância conjunta da variável REND nos três ambientes para avaliação dos híbridos intermediários. As variâncias residuais foram consideradas homogêneas quando a relação entre o maior e o menor quadrado médio residual não ultrapassou o valor sete (KRONKA, BANZATTO 1995). As médias obtidas por meio da análise de variância foram agrupadas pelo teste de Scott Knott a 5% de probabilidade pelo pacote ExpDes (Ferreira et al. 2013) do software R (R CORE TEAM, 2017).

Os resultados obtidos da eletroforese capilar foram transformados em uma matriz binária pelo software GeneMapper® v.4.1 (Applied Biosystems, Califórnia, USA). No programa Arlequin v. 3.5.2.2 (EXCOFFIER, LAVAL e SCHNEIDER, 2005) foi realizada a análise da variância molecular (AMOVA) para calcular a variabilidade genética entre e dentro dos genótipos e o índice de fixação alélica (F_{st}).

Mediante os resultados obtidos, foi gerado uma matriz de distância genética pelo pacote poppr do software R (KAMVAR, TABINA e GRUNWALD, 2014) baseado na distância de Nei & Li (1979), por se tratar de dados genéticos e binários. De acordo com a matriz gerada foi construída uma Neighbor-Net utilizando o programa SplitsTree (HUSON e BRYANT, 2006) para visualização o agrupamento entre os genótipos.

Para identificar a formação de grupos (K) dos genótipos e contribuição alélica dos genitores foi utilizado a análise bayesiana pelo programa BAPS -Bayesian Analysis of Population Structure (CORANDER et al. 2004).

Resultados e discussão

O resultado da análise conjunta dos ensaios de HPEs, nos três ambientes avaliados, apresentou coeficiente de variação (CV%) de 10,64%, revelando alta precisão experimental para ensaios conduzidos em campo (PIMENTEL GOMES, 2000).

Na análise conjunta, verificou-se significância ($P < 0,01$) nas fontes de variação Tratamentos, Ambientes e interação Tratamentos x Ambientes para a variável rendimento de grãos (REND), ou seja, os híbridos intermediários apresentaram respostas diferenciadas dentro de cada ambiente de teste (Tabela 2).

Tabela 2. Quadrados médios obtidos por análise de variância conjunta, coeficiente de variação e médias obtidas dos Híbridos Experimentais de Linhagens Parcialmente Endogâmicas (HPE) para a variável produtividade de grãos (REND) em kg. ha⁻¹ em três ambientes na safra 2016/2017.

FV	GL	REND
Bloco/Ambiente	3	1,54
Tratamentos	83	7,81**
Ambientes	2	173,37**
Trat x Amb	166	1,96**
Resíduo	249	1,55
CV(%)		10,64
Média Geral		11,69

ns,**: não significativo e significativo ao nível de 1%, pelo teste F.

Concordante com os resultados obtidos na análise conjunta, foi realizado o agrupamento de médias por Scott Knott (Tabela 3), incluindo as testemunhas nos três locais, a fim de verificar o padrão de ranqueamento dos híbridos. No agrupamento de médias, alguns híbridos avaliados apresentaram potencial produtivo similar ou superior às testemunhas comerciais (grupo a), dependendo das condições ambientais de cada ambiente avaliado.

Tabela 3. Agrupamento de médias por Scott-Knott a 5% de probabilidade dos vinte híbridos experimentais de linhagens parcialmente endogâmicas (HPE) mais produtivos em grãos (kg.ha⁻¹), nas localidades de Londrina, Santa Tereza do Oeste e Guarapuava, PR na safra 2016/2017.

Londrina		Santa Tereza do Oeste		Guarapuava	
220 x 116	15.642 aA	210 x 103	13.813 aA	220 x 112	15.262 aA
223 x 110	13.341 aA	204 x 112	13.034 aA	217 x 112	14.853 aA
209 x 103	14.505 aA	202 x 109	13.861 aA	204 x 109	14.082 aA
2B810	14.962 aA	218 x 119	12.902 aA	204 x 103	14.133 aA
210 x 103	12.964 aB	223 x 107	12.728 aA	203 x 113	15.468 aA
222 x 119	11.501 aB	220 x 116	14.469 aA	2B810	15.301 aA
203 x 111	10.940 aB	220 x 112	13.123 aA	222 x 119	12.655 aB
203 x 118	11.882 aB	204x116	14.507 aA	204 x 104	12.889 aB
218 x 119	11.803 aB	215 x 107	15.277 aA	223 x 112	12.321 aB
205 x 109	11.581 aB	210 x 113	13.641 aA	203 x 111	12.820 aB
215 x 101	10.792 aB	204 x 111	12.782 aA	203 x 118	12.283 aB
223 x 118	11.565 aB	205 x 109	12.782 aA	221 x 113	12.254 aB
222 x 101	10.915 aB	215 x 101	13.511 aA	223 x 107	13.114 aB
209 x 113	12.574 aB	215 x 104	13.860 aA	215 x 107	12.822 aB
220 x 103	11.253 aB	223 x 118	13.152 aA	204 x 111	12.908 aB
204 x 107	12.410 aB	220 x 103	13.298 aA	223 x 118	12.390 aB
218 x 101	11.542 aB	204 x 107	13.303 aA	204 x 107	13.117 aB
210 x 118	11.271 aB	207 x 109	12.651 aA	208 x 103	13.520 aB
219 x 101	11.803 aB	218 x 101	13.344 aA	210 x 118	12.795 aB
215 x 114	11.784 aB	214 x 107	13.648 aA	215 x 114	12.517 aB

*Letras minúsculas agrupamentos entre genótipos; letras maiúsculas agrupamento entre ambientes

Foram selecionados os vinte melhores HPE de cada localidade com base no primeiro quartil da amostra de médias, tendo em vista que um amplo número de HPEs foram agrupados em um mesmo conjunto dentro de cada ambiente de avaliação, permitindo observar um padrão de performance de determinados híbridos associados a parentais comuns que se repetiam ao longo dos ambientes, mesmo que inseridos em híbridos distintos. Dentre esses, os parentais machos (obtidos de PC0201) observados nos três ambientes foram 103, 107, 109, 111, 113, 118 e 119, enquanto os parentais fêmeas (obtidos do PC0202) identificados foram: 204, 215, 220 e 223.

Os demais HPEs apresentaram oscilações nas médias de produtividade, em função do ambiente de teste, o que era esperado uma vez que os componentes não aditivos para híbridos envolvendo cruzamento entre linhagens S_2 , estão em maior proporção que os componentes aditivos.

Analisando o desempenho dos HPEs, foi constatado que os híbridos 220x116 e 220x112 apresentaram elevado potencial produtivo, independente das variações ambientais. Já o híbrido 220x116 obteve alto potencial produtivo para ambientes específicos como Santa Tereza do Oeste e Londrina, porém em Guarapuava, Centro-Sul do Paraná, ambiente de elevada altitude e clima subtropical, a média de produtividade de grãos foi reduzida, sendo também observado nos materiais: 210x103, 218x119, 205x109, 215x101, 220x103 e 218x101.

Junior et al. (2017) observaram que híbridos intermediários extraídos de linhagens parcialmente endogâmicas (S_2) de milho tropical, apresentaram variabilidade e estabilidade dos materiais e rendimento de grãos semelhante e superior as testemunhas comerciais. Porém esse resultado não foi observado em todos os ambientes avaliados, concluindo que as diferentes condições climáticas exercem influência na variável rendimentos de grãos em híbridos intermediários.

Este padrão de resposta às oscilações ambientais são similares a trabalhos envolvendo híbridos simples, o que permite inferir que análises em estágios iniciais de endogamia tendem a otimizar o processo de seleção e/ou descarte de híbridos que não atendem os objetivos do programa de melhoramento, tais como seleção para adaptabilidade ampla ou específica a determinado conjunto de ambientes (OLIVOTO et al. 2017; PELEGRIN et al. 2017)

Buscou-se marcadores AFLP que permitissem separar grupos heteróticas das linhagens endogâmicas S_2 , conforme sua origem e analisar a variabilidade molecular intrínseca das duas populações e seus cruzamentos com as performances de produtividade em campo, a fim de acelerar o processo de seleção de genótipos superiores.

Foram obtidos 524 marcadores AFLP, dos quais 497 foram polimórficos, mostrando que estes híbridos intermediários apresentam alta variabilidade genética. A Análise de Variância Molecular (AMOVA) mostrou uma maior variabilidade dentro das populações (87,71%) do que entre os híbridos (12,29%). Isto mostra que se tem uma diferenciação genética de 12,29% entre as populações híbridas, fato confirmado pelo $F_{st}=0.12291$ (Tabela 4).

Tabela 4. Análise de Variância Molecular (AMOVA) das linhagens populações parcialmente endogâmicas (LPE) e híbridos experimentais de linhagens parcialmente endogâmicas (HPE).

Fonte de variação	GL	SQ	Componentes de variância	Porcentagem de variação
Entre populações	2	587,278	7,64000 Va	12,29
Dentro de populações	121	6596,843	54,51937 Vb	87,71
Total	123	7184,121	62,15936	
Índice de fixação (F_{st})	0,12291			

O grande número de fragmentos polimórficos obtidos pelo AFLP, mostra a eficiência dos marcadores em acessar as informações genéticas adquiridas em uma única corrida eletroforética (CAIXETA; FERRÃO e ZAMBOLIM, 2013), ressaltando a importância dos programas de melhoramento genético em ampliar a variabilidade genética dentro das linhagens provenientes de uma mesma base genética, a fim de obter híbridos com alto potencial produtivo, mesmo que parte dos locos (25%) não estejam fixos nos parentais S_2 .

Para analisar a divergência entre os genótipos de milho foi realizada uma análise bayesiana pelo software BAPS, onde ficou evidente que a ancestralidade de LPE0202 foi superior em relação a LPE0201 nos híbridos intermediários (Figura 1).

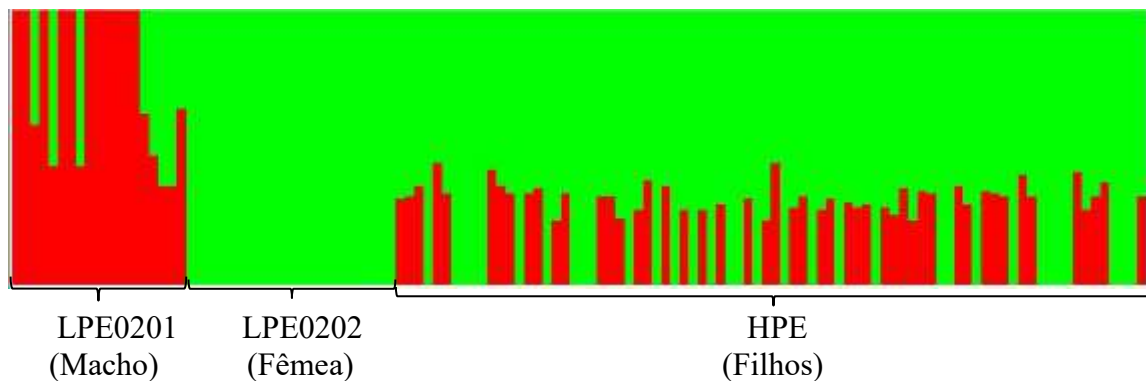


Figura 1. Porcentagem de ancestralidade dos parentais S_2 em relação aos híbridos intermediários de milho pelo software BAPS (Bayesian Analysis of Population Structure)

Foi realizada, ainda, a determinação de clusters (K), baseada apenas nas informações genéticas dos genótipos (Figura 2). Concordante com a Figura 1, apenas dois grupos foram

formados: LPE0201 e LPE0202 + HPE, onde os genótipos agrupados no mesmo cluster constitui da prevalência genética herdada. Foi observado que três materiais das linhagens LPE0201 se agruparam no cluster de LPE0202, isto é, ainda há variação na frequência genética das linhagens S₂ pois não estão em homozigose completa.

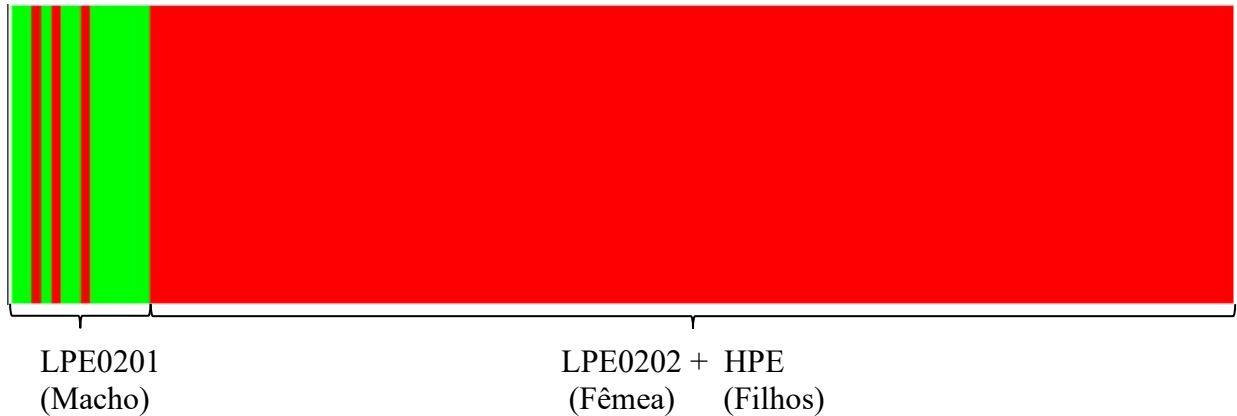


Figura 2. Determinação do número de clusters (K) nos genótipos de milho pelo software BAPS (Bayesian Analysis of Population Structure)

Uma hipótese para esse fato, possivelmente está ligado aos materiais base utilizados para composição das duas populações. Foram utilizados quatro híbridos simples comerciais distintos em cada população, sendo que, na população PC0202 dois híbridos são provenientes do portfólio de uma mesma empresa, ou seja, pertencem a um mesmo “*pool*” genético podendo ter parentais comuns, o que explica a predominância em determinados HPEs de uma estrutura populacional (Figura 1 e 2) similar a população PC0202 que pode ser reduzida se comparada a PC0201.

Deve ser destacado que um dos híbridos utilizados na composição da estrutura genética de PC0201 pertence a mesma empresa que detêm o registro de dois híbridos utilizados na composição da PC0202. Nesse sentido, é importante ressaltar que ao longo dos ciclos de recombinação para estabilização das populações, a seleção praticada pode tender a direcionar as frequências alélicas a um aumento ou redução para diferentes locus, dependendo do padrão agrônomico que se busca fixar na população estabilizada.

Baseado na distância de Nei & Li, foi realizado uma Neighbor-Net pelo software SplitsTree (Figura 3), a fim de observar a distribuição e divergência entre os genótipos para possíveis combinações heteróticas. A análise corrobora com os resultados apresentados pelo software BAPS, reforçando que a maior parte dos genitores femininos (LPE0202) se agrupam

geneticamente aos híbridos intermediários, havendo apenas a existência de um grupo divergente (cor azul) representado pela maioria das LPE0201.

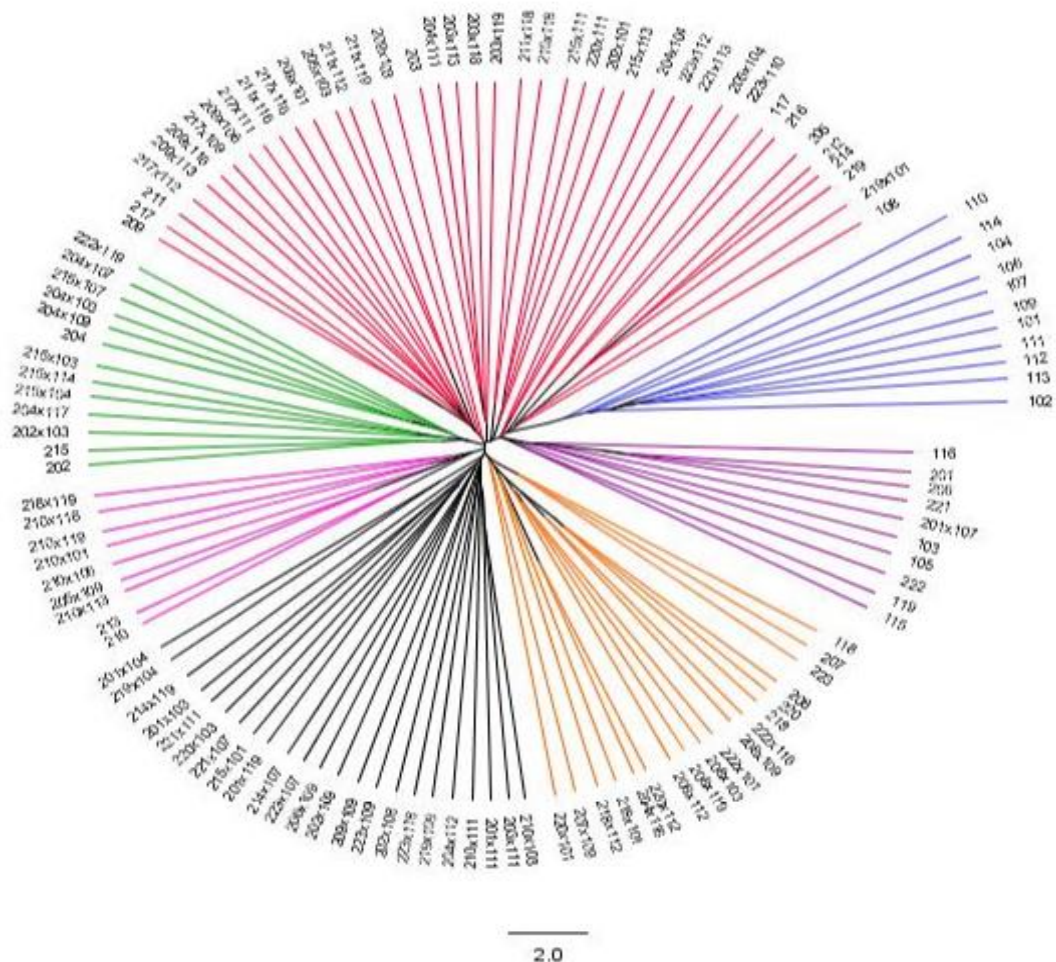


Figura 3. Neighbor-Net com base na distância de Nei & Li (1979), para todos os genótipos de milho (LPE0201, LPE0202 e HPE).

A maior parte dos híbridos com alto rendimento de grãos em campo são cruzamentos cujo genitores masculinos estão contidos justamente no grupo divergente em azul (101, 102, 104, 106, 107, 109, 110, 111, 112, 113 e 114), comprovando que linhagens contrastantes geneticamente, mesmo em estágio de endogamia parcial (S_2) resultam em híbridos de alta performance independentemente dos ambientes em que foram testados, permitindo inferir que os marcadores AFLP foram eficientes no agrupamento e classificação dos genótipos testados visando vigor híbrido.

Segundo Guimarães et al. (2007) e Lorencetti et al (2006) também verificaram correlação positiva entre rendimentos de grãos avaliados em campo e distância genética por meio do marcador molecular AFLP. Diversos tipos de marcadores moleculares foram avaliados

por Molin e colaboradores (2013) para estimar a diversidade genética em 48 variedades de milho. Entre os marcadores testados, o AFLP teve destaque por apresentar maior número de fragmentos polimórficos, obtendo assim maior número de informação genética e, conseqüentemente, melhor acurácia na delimitação dos grupos genéticos.

Além disso, deve ser destacado que o custo envolvido na utilização de AFLP é mais baixo se comparado a SSR ou SNPs, também é considerado um marcador robusto e confiável, além de ser rápido, por não precisar do conhecimento prévio da sequência requerida de DNA (FRASCAROLI et al. 2013; COSTA et al, 2016; RODRÍGUEZ-PEÑA et al. 2018; STONE et al. 2019).

Conclusões

- Os marcadores moleculares dominantes AFLP foram eficientes na identificação e clusterização dos padrões heteróticos mesmo em estágios de endogamia parcial (S_2).
- A conservação dos padrões heteróticos no estágio S_2 permite um descarte ou seleção eficiente de linhagens promissoras, para fases avançadas de endogamia.

4. ARTIGO B. CONSTRUÇÃO DE UMA REDE NEURAL MULTIPERCEPTRON APLICADA À CLASSIFICAÇÃO POR PRODUTIVIDADE DE GENÓTIPOS DE MILHO

Resumo

As Redes Neurais Artificiais são modelos de aprendizado de máquinas voltadas ao reconhecimento de padrões a partir de informações provenientes de diferentes origens (campo ou laboratório, por exemplo), destacando-se entre essas, as redes multilayer perceptron (MLP). No campo do melhoramento esta metodologia vem ganhando importância, uma vez que esta pode promover um incremento no ganho genético para variáveis de interesse, como o caso de produtividade de grãos na cultura do milho. Dessa forma, o presente trabalho desenvolveu uma arquitetura da rede MLP, com objetivo de desenvolver e testar um modelo de classificação voltado à seleção e descarte de genótipos de milho (linhagens e híbridos com endogamia parcial – S₂) com base no rendimento de grãos, utilizando conjuntos de dados meteorológicos de três ambientes no Estado do Paraná e marcadores moleculares dominantes. A arquitetura de rede neural construída foi eficiente na classificação de genótipos por produtividade e a utilização conjunta de dados de marcadores moleculares e meteorológicos promoveram o aumento da capacidade preditiva do modelo.

Palavras-chave: *Zea mays* L.; Marcadores AFLP; Redes Neurais Artificiais

Abstract

Artificial Neural Networks are models of machine learning aimed at pattern recognition based on information from different sources (field or laboratory, for example), standing out among these, multilayer perceptron networks (MLP). In the field of breeding, this methodology has been gaining importance, since it can promote an increase in genetic gain for variables of interest, such as the case of grain productivity in maize. Thus, the present work developed an MLP network architecture, with the objective of developing and testing an efficient classification model focused on the selection and disposal of maize genotypes (lines and hybrids with partial inbreeding - S₂) based on grain yield, using meteorological data sets from three environments in the State of Paraná and dominant molecular markers. The built-in neural network architecture was efficient in classifying genotypes by productivity and the combined use of marker and meteorological data promoted an increase in the predictive capacity of the model.

Key-words: *Zea mays* L.; AFLP marker; Artificial Neural Networks

Introdução

A Inteligência Artificial (IA) tem ganhado destaque em diversas áreas por ser uma ferramenta computacional que através de algoritmos indica a melhor decisão em uma determinada situação, semelhante a habilidade cognitiva do ser humano. Diversos ramos utilizam esse método e garante eficiência, como na área médica (OLIVEIRA, 2019), industrial (RAHMANIFARD; PLAKSINA, 2019), educação (SHIN, 2020), ambiental (KAAB et al. 2019), agricultura (ELI-CHUKWU, 2019) entre outros.

Esse método consiste em treinar uma rede com experiências acumuladas e torná-lo automatizado. (JHA et al. 2019). As Redes Neurais Artificiais (RNAs) são modelos de aprendizado de máquinas que reconhecem padrões e apresentam retropropagação combinada, através de neurônios artificiais interconectados (SHANMUGANATHAN, 2016). Ainda, dentro das arquiteturas disponíveis das RNAs, a multilayer perceptron (MLP) busca ajustar os pesos entre os neurônios (retropropagação de erro) envolvendo mais de uma camada oculta dentro da rede, possibilitando um aumento na acurácia preditiva de características de interesse, mesmo quando distribuições de probabilidade são “*a priori*” desconhecidas (DEB; OGUNFUNMI, 2017).

As MLPs, assim como outras ferramentas computacionais têm sido utilizadas com maior frequência na agricultura, denominada agricultura inteligente ou *smart farm* (PIVOTO et al. 2017). Essas ferramentas estão presentes em diversos processos agrícolas que vão desde características de interesse do fenótipo, identificação de doenças, qualidade nutricional, previsões agrometeorológicas, interação genótipo e ambiente (GxA), genotipagem entre outros. Em todos os processos citados, os dados e informações gerados são armazenados e utilizados para treinamento da rede a fim de obter uma classificação e seleção eficiente futura (PATRÍCIO; RIEDER, 2018).

Assim, recentemente programas de melhoramento de plantas têm apostado no uso de MLP para estimar valores genéticos (SILVA et al. 2016), avaliar diversidade genética (JIANG, 2020), identificar e selecionar genótipos de alta performance (PARMLEY et al. 2019) e também para seleção genômica (GONZÁLEZ-CAMACHO et al. 2018). Azodi e colaboradores (2019) testaram doze tipos de algoritmos aplicados em seis espécies de plantas juntamente com dados genotípicos e fenotípicos onde foi observado que a acurácia de predição depende da característica avaliada e, quanto maior a informação, maior a eficiência, além de destacar que as RNAs são superiores podendo ser utilizadas para solucionar problemas adversos da seleção genômica.

Na cultura do milho, a predição e classificação de genótipos com maior rendimento de grãos são fundamentais para o êxito de um programa de melhoramento genético, no entanto, conforme destacam Meena et al. (2017), para alcançar a eficiência nesse processo, a identificação de grupos heteróticos é primordial para maximizar o vigor híbrido. Lana et al. (2017) citam que além dos dados genéticos as análises de dados climatológicos são relevantes no comportamento fenotípico da cultura do milho, principalmente precipitação e temperatura do ar, fatores estes, fundamentais para um bom desenvolvimento da planta e produtividade. Dessa forma, algumas empresas públicas e privadas de melhoramento genético tem recorrido à ferramenta de MLP para obter uma direção acurada na seleção de materiais produtivos, além de observar parâmetros da planta em relação as oscilações ambientais.

Diversos estudos apontam que o treinamento supervisionado maximiza o ganho genético e a predição de uma determinada cultura, quando se utiliza um maior conjunto de dados, sejam eles, características quantitativas e/ou qualitativas, marcadores moleculares, imagens, sensores e até dados agrometeorológicos. (FALK et al. 2020; PARMLEY et al. 2019; HESLOT et al. 2012). Todas essas características podem ser incluídas no MLP, motivo esse que diferencia dos outros métodos, proporcionando um incremento na classificação.

Dessa forma, o presente estudo tem o objetivo de obter/construir um padrão de classificação para seleção ou descarte de genótipos de milho com foco no rendimento de grãos por meio de um modelo de redes neurais multicamadas treinadas através de dados conjuntos de marcadores moleculares dominantes e dados agrometeorológicos em diferentes ambientes.

Material e Métodos

Material Genético e Experimento em campo

No Programa de Melhoramento de Milho do Instituto de Desenvolvimento Rural do Paraná - IAPAR-EMATER (IDR-Paraná) foram sintetizadas duas populações de milho experimental (PC0201 e PC0202) e intercruzados seguindo um modelo de dialelo completo na safra 2005/2006. Foram obtidas 90 linhagens S₂ de cada população e submetidas a um cruzamento dialelico parcial circulante interpopulacional (CDPCI), resultando em 360 híbridos.

Esses híbridos intermediários e seus parentais foram avaliados na safra 2012/2013 e selecionadas os materiais com maior produção de grãos e menor índice de acamamento de plantas, obtendo 42 linhagens parcialmente endogâmicas (LPE) e 82 híbridos experimentais de linhagens parcialmente endogâmicas (HPE). Os HPE foram avaliados em três locais no Estado do Paraná (Londrina, Santa Tereza do Oeste e Guarapuava) com duas testemunhas comerciais

(DKB210 e 2B810). Já as LPE foram avaliadas apenas em Londrina e Santa Tereza do Oeste com uma testemunha comercial (2B810) mais as populações base (PC0201 e PC0202), os ensaios dos HPE e LPE foram avaliados na safra 2016/2017 (Tabela 1).

Tabela 1. Descrição dos ambientes de avaliação dos 82 híbridos intermediários, 19 linhagens S₂ da população PC0201 e 23 linhagens S₂ da população PC0202, no Estado do Paraná. Safra 2016/2017.

Ambiente (Código)	Altitude	Lat	Long	* Clima	Solo
Londrina (LD)	569 m	23°22'0.06"S	51° 9'55.61"O	** Cfa	Latossolo Vermelho distrófico
Santa Tereza do Oeste (STO)	749 m	25° 08'S	53° 58'O	*** Cfa	Latossolo Vermelho distroférico típico
Guarapuava (GUA)	1038 m	25°23'51.87"S	51°32'35.25"O	*** Cfb	Latossolo Bruno Álico e associações com Cambissolo Álico

* Segundo o sistema de classificação climática de Köppen-Geiger. ** Cfa: clima subtropical úmido. *** Cfb: clima subtropical úmido mesotérmico

Os experimentos foram conduzidos em delineamento em blocos casualizados com duas repetições por local, onde cada parcela experimental foi constituída de 5 plantas por metro linear após o desbaste, sendo que, a cada parcela experimental foi constituída por uma linha de 5 metros de comprimento e 0,80 m entre linhas. Foi analisado o rendimento de grãos (REND) de cada parcela, no qual as espigas foram colhidas e os grãos obtidos de cada parcela foram pesados, ajustados para kg.ha⁻¹ e corrigidos para uma umidade padrão de 14,5%.

Os valores obtidos de rendimento de grãos foram divididos em duas classes:

Classe 1 – genótipos com médias de produção de grãos superiores à mediana, em valores absolutos, do conjunto de dados dentro de cada local e grupo genotípico (híbridos ou linhagens).

Classe 2 – genótipos com média de produção de grãos abaixo da mediana dos dados, dentro de cada local e grupo genotípico (híbridos ou linhagens).

Extração DNA e Análise do Marcador Molecular – AFLP

Para a extração de DNA foi utilizado o procedimento proposto por Doyle e Doyle (1997) e foram coletadas cinco folhas de cada genótipo (82 HPE e 42 LPE) dez dias após emergência em campo. Aproximadamente 700 ng do DNA genômico foram digeridos utilizando enzimas de restrição de corte frequente MseI (1U) e corte raro EcoRI (5U) e em seguida, os fragmentos

gerados foram ligados a adaptadores específicos.

Na amplificação pré-seletiva foi utilizado o kit GoTaq® Green Master Mix (Promega, Winchester, EUA), já para a amplificação seletiva foram utilizadas quatro combinações de primers marcados com fluoróforos EcoRI ACG (FAM-azul) / MseI-CAG; EcoRI AGG (PET-vermelho) / MseI-CTG; EcoRI AAG (NED-amarelo) / MseI CTG e EcoRI AAG (HEX-verde) / MseI-CTC).

Posteriormente foi feito um sistema multiplex com a diluição de 2,0 µL das reações seletivas com 3 µL de água ultrapura, para resolução dos produtos de PCR, e a cada 1,0 µL dessa mistura foi adicionado 0,2 µL de LIZ 500 e 8,8 µL de formamida Hi-Di Pop-7. As reações foram submetidas a eletroforese capilar (ABI 3500XL -Applied Biosystems, Califórnia, EUA) e os resultados foram gerados em uma matriz binária utilizando o software GeneMapper v.4.1 (Applied Biosystems, Califórnia, EUA).

Dados Agrometeorológicos

Os dados agrometeorológicos utilizados foram as temperaturas máxima, média e mínima e precipitação diárias disponibilizados pelo Sistema de Tecnologia e Monitoramento Ambiental do Paraná – SIMEPAR no período referente aos ensaios de campo em cada ambiente. Para os dados de temperatura foram utilizados padrões definidos conforme a literatura para a cultura do milho, nos quais temperaturas noturnas máxima e mínima são definidas em 24°C e 12°C respectivamente e nas temperaturas diurnas 30°C a máxima e 19°C a mínima (FANCELLI E DOURADO-NETO, 2000). Também foi considerado a faixa de temperatura ideal para a cultura, que se encontra entre 15°C a 26°C, sendo que abaixo do mínimo estabelecido, as fases de desenvolvimento se prolongam devido à redução dos processos metabólicos, e acima de 26°C ocorre o processo inverso, ou seja, as etapas de desenvolvimento são aceleradas.

As temperaturas médias obtidas não apresentaram grandes oscilações, portanto, foram utilizados classificadores baseados nos picos das temperaturas máximas e mínimas ($T_{máx}$ e $T_{mín}$) em relação aos períodos críticos e não críticos de desenvolvimento da cultura do milho, conforme a Tabela 2:

Tabela 2. Classificadores dos períodos não críticos e críticos da cultura do milho baseados nos picos de temperaturas máximo e mínimo.

Períodos não críticos:	Períodos críticos:
$T_{máx} > 30^{\circ}\text{C}$ e $T_{mín} < 10^{\circ}\text{C}$ = classe 0	$T_{máx} > 26^{\circ}\text{C}$ e $T_{mín} < 15^{\circ}\text{C}$ = classe 0
$T_{máx} \leq 30^{\circ}\text{C}$ e $T_{mín} > 10^{\circ}\text{C}$ = classe 1	$T_{máx} \leq 26^{\circ}\text{C}$ e $T_{mín} > 15^{\circ}\text{C}$ = classe 1

(Continuação)

Períodos não críticos:	Períodos críticos:
$T_{\text{máx}} \leq 30^{\circ}\text{C}$ e $T_{\text{mín}} < 10^{\circ}\text{C}$ = classe 2	$T_{\text{máx}} \leq 26^{\circ}\text{C}$ e $T_{\text{mín}} < 15^{\circ}\text{C}$ = classe 2
$T_{\text{máx}} > 30^{\circ}\text{C}$ e $T_{\text{mín}} > 10^{\circ}\text{C}$ = classe 3	$T_{\text{máx}} > 26^{\circ}\text{C}$ e $T_{\text{mín}} > 15^{\circ}\text{C}$ = classe 3

As variáveis selecionadas para o grupo de temperatura (TEMP), foram divididas em 10, 20, 30 e 40 dias após o florescimento (temp_Ma10X; temp_Ma20X; temp_Ma30X; temp_Ma40X) respectivamente, e antes do florescimento (temp_Men10X; temp_Men20X; temp_Men30X; temp_Men40X).

A precipitação foi outro fator climático avaliado considerado que para a planta de milho desenvolver cada etapa do seu ciclo, é necessária uma demanda hídrica de 400mm a 600mm. No início do desenvolvimento da cultura as exigências pluviométricas são baixas e, conforme há um aumento no desenvolvimento da planta, há também uma maior exigência para água, sendo que, a fase entre 15 dias antes e após o pendoamento é considerada crítica, pois o déficit hídrico nesse período ocasiona perdas representativas de produtividade (BERGARMASCHI et al. 2004).

Dessa forma, o volume pluviométrico acumulado em cada local foi mensurado com intervalos de 10 dias para períodos entre 10 a 40 dias após o florescimento (chuva_Ma10X; chuva_Ma20X; chuva_Ma30X; chuva_Ma40X) e também antes do florescimento (chuva_Men10X; chuva_Men20X; chuva_Men30X; chuva_Men40X).

Diante do exposto, para eficácia e compreensão das diferenças em função da temperatura e precipitação foi aplicado como fase crítica o período de 10 e 20 dias antes e após o florescimento. As precipitações nas fases não críticas foram baseadas em intervalos de 10mm e para as fases críticas os intervalos foram de 25mm (Tabela 3).

Tabela 3. Classificação de precipitação em milímetros (mm) das fases não críticas e críticas.

Fases não críticas	Fases críticas
$\geq 20\text{mm} = 2$	$\geq 50\text{mm} = 2$
$10 - 20\text{mm} = 1$	$25-50\text{mm} = 1$
$\leq 10\text{mm} = 0$	$\leq 25\text{mm} = 0$

Análises estatísticas de campo e marcadores

Foram realizadas análises de variância conjuntas através do software R (R Core Team, 2017) para a variável rendimento de grãos (REND), nos ambientes avaliados das LPE e HPE.

Os resultados genotípicos dos marcadores moleculares AFLP foram obtidos pela eletroforese capilar e transformados em uma matriz binária pelo software GeneMapper® v.4.1 (Applied Biosystems, Califórnia, USA), e pontuada quanto a presença (1) ou ausência (0) de bandas polimórficas, sendo que, os marcadores monomórficos foram retirados da análise.

De modo a analisar a eficiência dos marcadores, foi realizada uma análise classificatória prévia (ACPRV) utilizando como dados de entrada a matriz binária dos marcadores na rede neural. Os dados de saída foram as identificações relativas a híbridos intermediários (Classe 3), linhagens provenientes da população PC0201 (Classe 1) e linhagens obtidas da população PC0202 (Classe 2).

Redes Multilayer Perceptron (MLP)

A arquitetura da rede utilizada para esta análise prévia foi 636 neurônios na camada de entrada, referente a cada marcador molecular, 3 camadas ocultas com 100 neurônios por camada e 3 neurônios na camada de saída, referente às 3 classes genotípicas a serem preditas, sendo utilizadas 5000 épocas para ajuste do modelo, utilizando o pacote H₂O (LEDELL; POIRIER, 2020) no software R. (R CORE TEAM, 2017).

Foi utilizado nas amostras, 80% dos dados para treinamento e 20% para validação, sendo este processo repetido 100 vezes, onde em cada repetição, os pesos e viés eram armazenados, bem como a matriz de confusão para determinação da acurácia do modelo.

Foi construído um gráfico usando Microsoft Excel para analisar a importância de cada uma das variáveis moleculares e meteorológicas, inserida na rede MLP para Análise Classificatória Supervisionada (ACS). Após a análise de importância das variáveis, foi efetuada uma montagem de Boxplot em relação a porcentagem de acurácia classificatória através da MLP para as variáveis: moleculares e meteorológicas separadas e uma agrupada para os dois conjuntos simultaneamente. Em seguida foi montado um heatmap com base nos valores de importância das variáveis dentro dos conjuntos de variáveis similares.

A rede neural utilizada para ACS, foi composta por uma camada de entrada na qual os neurônios continham as informações dos marcadores, através de uma matriz binária, mais os dados meteorológicos convertidos para uma distribuição discreta. O volume inicial de informações de entrada possuía mais de 500 neurônios que posteriormente teve este número reduzido através de uma análise de importância das variáveis, chegando ao número final de 60 neurônios na camada de entrada. Para as demais camadas, foram utilizadas 3 camadas ocultas contendo 100 neurônios cada e uma camada de saída com 2 neurônios correspondentes a identificação de genótipo selecionado (genótipo 1) ou descartado (genótipo 2). Para o

treinamento foi utilizada uma relação padrão de 80/20, ou seja, 80% do conjunto inicial de dados para treinamento e 20% para validação.

Para o treinamento da rede foram utilizadas 5000 épocas ou interações com backpropagation, ou seja, é realizado uma correção dos pesos sinápticos onde resposta real da rede é subtraída da desejada, dessa forma com a subtração, um sinal de erro é criado e este é propagado para trás, difundindo pela rede de forma contrária, por isso, retropropagação. Nesse caso os pesos sinápticos são ajustados para que o conjunto de saída seja o mais próximo, estatisticamente, do desejado.

Como mecanismo para evitar overfitting, foi adotado a metodologia de early stopping, ou seja, é calculada a precisão da classificação nos dados de validação no final de cada época. Quando a precisão da classificação nos dados de validação apresenta saturação, o sistema é parado imediatamente evitando overfitting.

Resultados e Discussão

Através das ANOVAS individuais foi verificado que as relações entre o maior e menor QM residual foi de 2,62 para linhagens e 1,72 para híbridos, indicando que as análises conjuntas podem ser conduzidas sem a necessidade de descarte de ambientes. Através das ANOVAS conjuntas das LPE e HPE foram verificados coeficientes de variação de 14,66 e 10,64 respectivamente, indicando boa precisão experimental em ambos os conjuntos de genótipos.

A interação genótipo com ambiente (GxA) foi considerada significativa ($p < 0,05$) para HPE, indicando que houve um comportamento diferencial dos genótipos avaliados ao longo dos ambientes de teste. Nesse caso os processos de classificação das produtividades por meio de redes neurais, deve considerar como parâmetro de entrada em um dado neurônio da camada inicial, o efeito de ambientes, de modo que o ajuste do modelo seja balizado considerando as variações ambientais.

A não significância para GxA relativa às linhagens, indicam que os dados médios poderiam ser utilizados no processo de classificação. No entanto, considerou-se que o fato das redes neurais utilizarem o vetor de identificação ambientes como atributo classificatório, e tendo em vista o efeito significativo da fonte de variação ambientes, tal fato indica que o uso de dados das linhagens separados por local, tende a agregar positivamente no treinamento do modelo (Tabela 4).

Tabela 4. Quadrados médios obtidos por análise de variância conjunta, coeficiente de variação e médias obtidas das Linhagens Parcialmente Endogâmicas (LPE) e Híbridos Experimentais de

Linhagens Parcialmente Endogâmicas (HPE) para a variável produtividade de grãos (REND) em kg. ha⁻¹ na safra 2016/2017.

FV	LPE		HPE	
	GL	REND	GL	REND
Bloco/Ambiente	2	1,35	3	1,54
Genótipos	21	23,01**	83	7,81**
Ambientes	1	102,28**	2	173,37**
GxA	21	1,84	166	1,96**
Resíduo	42	2,01	249	1,55
CV (%)		14,66		10,64
Média Geral		8,47		11,69

ns,**: não significativo e significativo ao nível de 1%, pelo teste F, respectivamente

De modo a se testar a eficiência dos indicadores da rede em delinear informações e classificar padrões, unicamente com base nos dados dos marcadores moleculares, foi utilizada a mesma arquitetura da rede neural com 50 repetições (Rp) para predição de classes genótípicas e padrões heteróticos, sendo estes constituídos pelas linhagens provenientes da população PC0201 (classe 1), linhagens de PC0202 (classe 2) e híbridos intermediários (classe 3) resultantes do inter cruzamento das linhagens.

A partir dos dados da matriz de confusão dentro de cada repetição foi possível constatar uma capacidade preditiva média de 89%, dentre as quais, 3 repetições apresentaram pico máximo de 100% de capacidade preditiva na determinação das classes genótípicas, 23 repetições apresentaram média de capacidade preditiva acima de 90%, enquanto as 74 repetições restantes apresentaram valores próximos ou acima de 80%, demonstrando que os marcadores possuem informações relevantes para o uso em classificações associadas à base genética dos materiais testados, mesmo que parte dos locos não estejam fixos nos parentais S₂ (Figura 1).

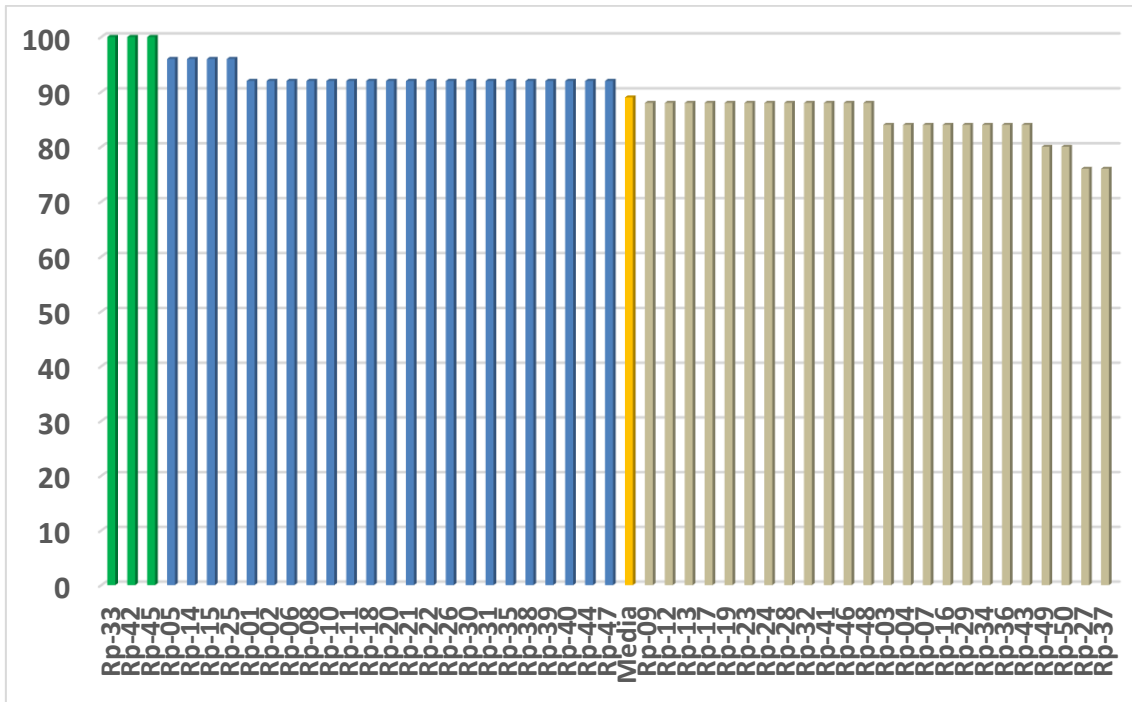


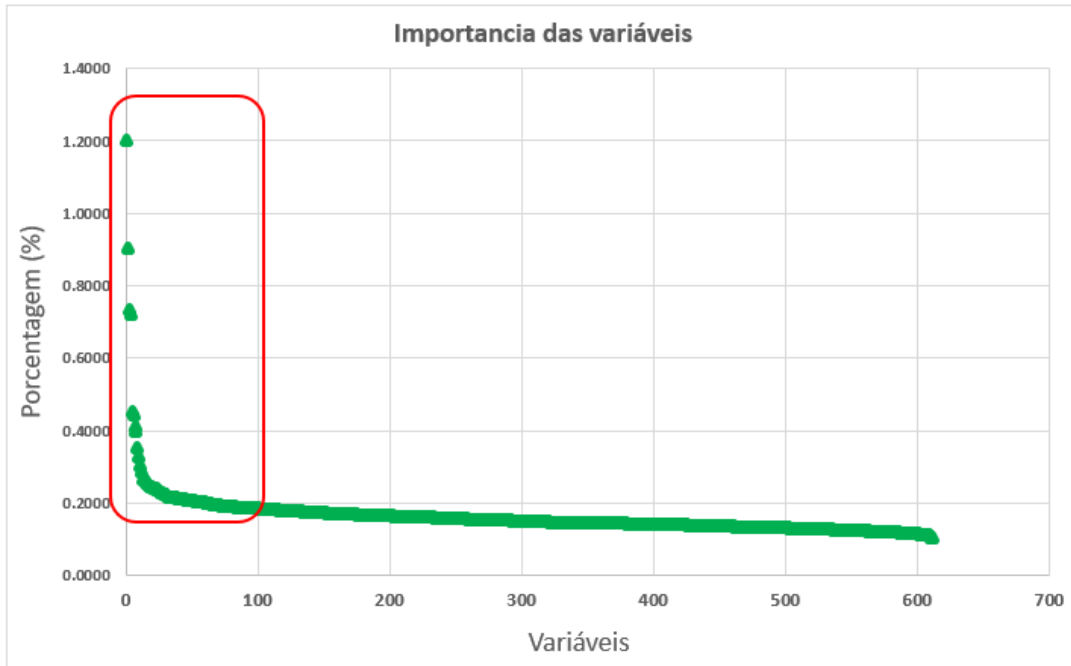
Figura 1. Percentagem de capacidade preditiva das classes genóticas e padrões heteróticos em 50 repetições com reamostragem da rede neural.

Segundo Balestre et al. (2008) e Bernal, Guerra e Oliveira (2017) a identificação de grupos ou padrões heteróticos é a chave do sucesso em programas de melhoramento de milho, uma vez que a precisão na escolha das linhagens maximiza o vigor híbrido. Wani et al. (2017) ressaltam que o conhecimento prévio dos genótipos evita o descarte errôneo e avaliações desnecessárias em campo, além de maximizar a precisão dos grupos heteróticos uma vez que nos cruzamentos manuais pode ocorrer contaminação genética.

Para otimizar o processo de classificação, foi realizada uma análise de importância relativa das variáveis, dentre as quais, foram incluídas a matriz binária dos marcadores moleculares AFLP, os dados pluviométricos e de temperatura convertidos para uma distribuição discreta, assim como um vetor comum que indica as codificações de ambiente.

A partir de um ponto de corte acima de 0,2000% (Figura 2a), onde se detectou uma relação de estabilização e início de decréscimo das variáveis, foram selecionadas 60 variáveis para a continuidade das análises de classificação de produtividade, onde essas foram agrupadas em 4 conjuntos (Figura 2b), sendo estes um conjunto relativo aos marcadores (marc), conjunto relativo aos dados de temperatura (temp), conjunto relativo aos dados pluviométricos (pluvi) e um conjunto para a identificação de ambientes (amb).

(a)



(b)



*Ma: acumulado de dias após florescimento; Men: número de dias antes do florescimento

Figura 2. Análise das variáveis moleculares e meteorológicas **(a)** importância (%) das variáveis em relação à classificação de produtividade. **(b)** Distribuição da importância das variáveis apresentada por cores azuis, sendo que, azul claro indicam maior importância e azul escuro menor, em relação à classificação rendimento de grãos.

Conforme, Gedil e Menkir (2019) observaram a utilização conjunta das variáveis de ambientes associadas à informação de marcadores acelera o ganho genético promovendo uma substituição ágil dos materiais presentes nos programas de melhoramento.

É possível constatar ainda pela figura 2b, que o ambiente e as variáveis temperaturas com 20 e 30 dias antes do florescimento são fatores fundamentais para obter genótipos produtivos. Este fato está associado em que o fenótipo varia conforme as condições climáticas do ambiente em teste e o período da temperatura que apresentam maiores percentuais de explicação para o conjunto temp mostra-se importante para o estabelecimento das primeiras folhas nessas fases (Ritchie, Hanway e Benson, 1993). Os autores ainda destacam em seu trabalho que períodos entre 15 dias antes e 15 dias após o florescimento seriam críticos para a cultura de milho, uma vez que neste intervalo, de acordo com Nielsen (2005), a cultura necessita de condições térmicas e hídricas suficientemente necessárias para a formação de órgãos reprodutores masculino e feminino, assim como a viabilidade desses na polinização, fecundação e fases iniciais de enchimento de grãos.

Considerando as variáveis associadas ao volume pluviométrico com 10 dias (Chuva_Ma10X) e 20 dias (Chuva_Ma20X) após o florescimento, Nielsen (2005) considera que o déficit hídrico nessas fases pode causar reduções significativas no rendimento final, uma vez que a planta encontra-se em pleno enchimento de grãos e absorção de nutrientes para composição final do grãos. Bergamaschi et al. (2004), através de modelo de regressão, também constataram que a maior redução no rendimento de grãos do milho ocorre devido ao déficit hídrico no período compreendido entre a polinização e o desenvolvimento inicial do grão.

Após a definição das variáveis finais para uso no modelo de classificação foi gerado um *looping* com 50 repetições considerando os parâmetros associados à arquitetura de rede proposta. Para três sistemas de entrada utilizados (Figura 3a) foi constatado que o modelo envolvendo os dados meteorológicos e matriz de marcadores selecionados apresentou os maiores percentuais de capacidade preditiva média (61%) com repetições atingindo picos de 72%, em que destas foram armazenadas em matrizes de pesos e vieses.

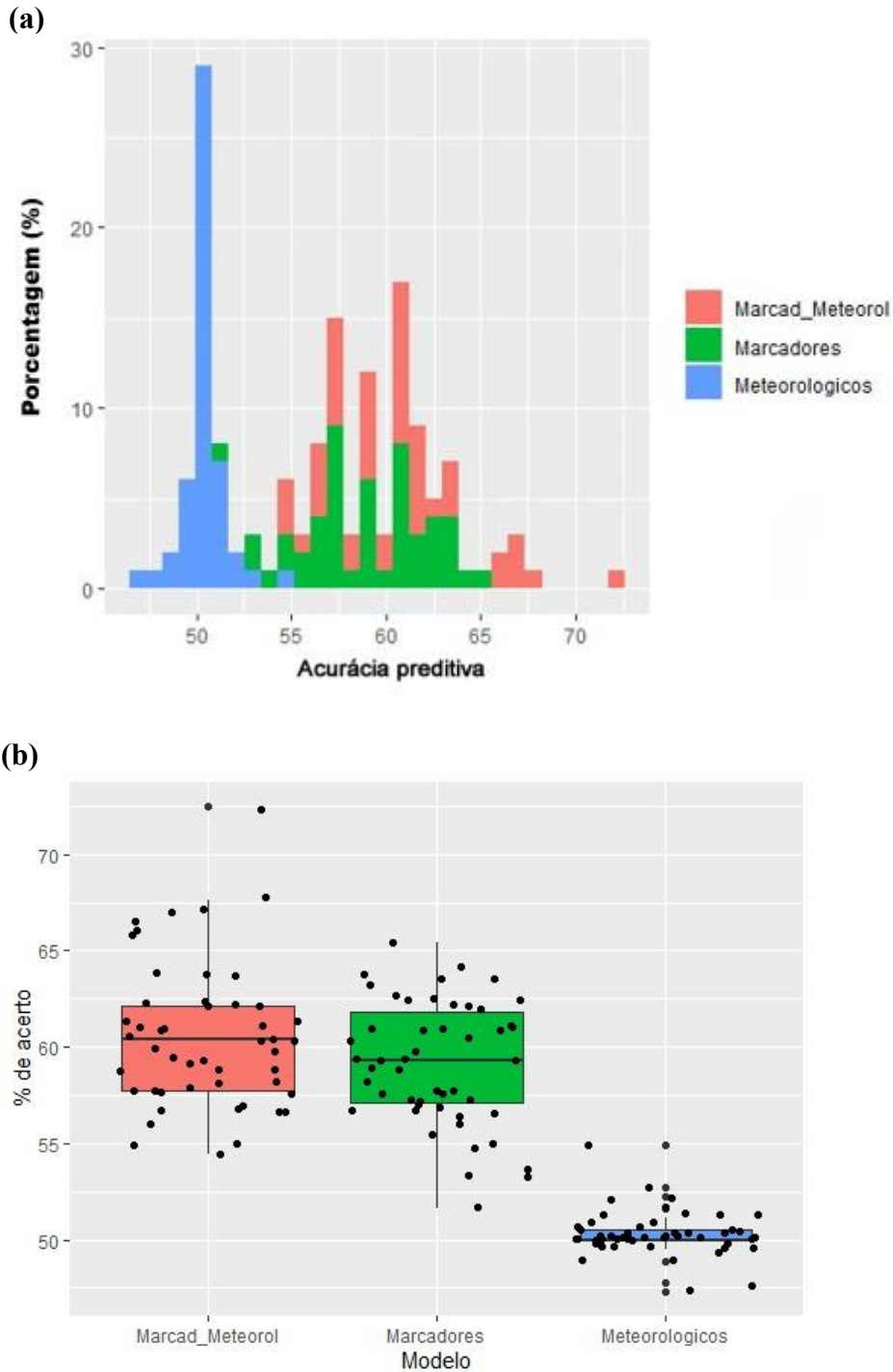


Figura 3. Dados de entrada: moleculares com agrometeorológico, marcadores moleculares e agrometeorológico (a) Distribuição das variáveis em barra, em relação a porcentagem de acerto através da MLP (b) BoxPlot das variáveis em relação a porcentagem de acerto pela rede MLP.

Para os dois modelos que consideraram somente os dados de marcadores e somente os dados meteorológicos, foram verificadas capacidade preditiva de 59% e 50%, respectivamente, onde o maior pico registrado foi para o modelo envolvendo somente marcadores com 65% de acurácia, estando todos abaixo do modelo que considerou na camada de entrada da rede neural, os dados de marcadores e meteorológicos conjuntamente.

Segundo Nehe et al. (2019), como a interação GxA é responsável diretamente pela variação fenotípica, está também associada aos ganhos genéticos. Utilizando esse princípio associado as redes neurais, Azevedo et al. (2015) realizaram a seleção indireta em genótipos de alface considerando a influência da interação GxA no período de florescimento, característica de alta importância comercial, onde obtiveram ganhos de seleção de 7,02% a 9,30%, utilizando uma rede neural treinada com dados de dois experimentos. Qiao et al. (2019) constataram que as análises integradas de respostas genômicas e diferentes ambientes, são poderosas para revelar uma arquitetura genética subjacente, que pode variar em tamanho e direção, considerando efeitos contínuos de genes individuais, ao lado de um gradiente ambiental.

Segundo Parmley et al. (2019) a interação de esforços multidisciplinares, entre estes, técnicas de M.L., seleção genômica e informações ambientais, podem melhorar a informação ao melhorista sobre o potencial genético dos genótipos candidatos, buscando um posicionamento e seleção otimizado para diferentes sistemas de manejo ou ambientes de cultivo.

Nesse sentido, no presente estudo, com base nos valores obtidos de capacidade preditiva classificatória, a partir das variáveis meteorológicas e de marcadores AFLP, foi possível obter um modelo de seleção com ênfase em classes de produtividade em milho com níveis intermediários de endogamia (S_2), tanto para linhagens quanto para híbridos.

Conclusões

- A arquitetura de rede neural construída foi eficiente na classificação de genótipos por produtividade
- A utilização conjunta de dados de marcadores moleculares e meteorológicos promoveram o aumento da capacidade preditiva do modelo.

5. CONSIDERAÇÕES GERAIS

A metodologia aplicada para identificação e clusterização dos padrões heteróticos em linhagens parcialmente endogâmicas (S_2) por meio de marcadores moleculares dominantes AFLP, foram eficientes mesmo quando oriundas de um mesmo “*pool*” genético comum, permitindo a seleção correta de linhagens promissoras para formação de híbridos intermediários.

A construção de uma arquitetura de rede neural Multilayer Perceptron, utilizando como base os valores obtidos de capacidade preditiva classificatória, a partir das variáveis meteorológicas e de marcadores AFLP, também foi eficiente na classificação de genótipos com alto rendimento de grãos. As variáveis utilizadas para modelagem da rede computacional promoveram um aumento da capacidade preditiva do modelo, mesmo quando as linhagens e híbridos apresentarem níveis intermediários de endogamia (S_2).

Visto isso, estratégias para identificação de padrões heteróticos como marcadores AFLP e Multilayer Perceptron são eficazes e econômicos para predição e seleção de híbridos intermediários de alta performance produtiva, podendo ser utilizados em programas de melhoramento genético de milho privados e públicos.

REFERÊNCIAS

- ABEL, F. **Análise textual automática: Apreensibilidade e qualidade da informação na área da saúde**. 2016. Disponível em: <https://www.researchgate.net/figure/Figura-10-Representacao-de-uma-Rede-Neural-ou-Multilayer-Perceptron_fig3_313063807>. Acesso: 04/05/2020.
- ABRISHAMI, N.; SEPASKHAH, A.; SHAHROKHANIA, M. H. Estimating wheat and maize daily evapotranspiration using artificial neural network. **Theoretical and Applied Climatology**, v. 135, n. 3-4, p. 945-958, 2019.
- AHMAD, F.; AKRAM, A.; FARMAN, K.; ABBAS, T.; BIBI, A.; KHALID, S.; WASEEM, M. Molecular Markers and Marker Assisted Plant Breeding. Current Status and their Applications in Agricultural Development. **J Environ Agric Sci**, v. 11, p. 35-50, 2017.
- ALAMEER, Z.; ABD ELAZIZ, M.; EWESS, A. A.; YE, H.; JIANHUA, Z. Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. **Resources Policy**, v. 61, p. 250-260, 2019.
- AMOM, T.; NONGDAM, P. The use of molecular marker methods in plants: a review. **International Journal of Current Research and Review**, v. 9, n. 17, p. 1-7, 2017.
- AN, J.; LI, W.; LI, M.; CUI, S.; YUE, H. Identification and classification of maize drought stress using deep convolutional neural network. **Symmetry**, v. 11, n. 2, p. 256-262, 2019.
- ANDRADE, L. R. H. Vulnerabilidade genética e relação entre os germoplasmas comerciais de milho no Brasil. In: **Congresso Brasileiro de Plantas**. Anais. Goiânia, 2015.
- ARAÚJO, P. M.; PATERNIANI, E. **Melhoramento de populações**. In: DESTRO, D.; MONTALVÁN, R. **Melhoramento Genético de Plantas**. Londrina: EDUEL, p.311-330, 1999.
- ARTUZO, F. D.; FOGUESATTO, C. R.; MACHADO, J. A. D.; DE OLIVEIRA, L.; DE SOUZA, Â. R. L. O potencial produtivo brasileiro: uma análise histórica da produção de milho. **Revista em Agronegócio e Meio Ambiente**, v. 12, n. 2, p. 515-540, 2019.
- AZODI, C. B., BOLGER, E., MCCARREN, A., ROANTREE, M., DE LOS CAMPOS, G., SHIU, S. H. Benchmarking parametric and Machine Learning models for genomic prediction of complex traits. **G3: Genes, Genomes, Genetics**, v. 9, n. 11, p. 3691-3702, 2019.
- BARRETO, R.; SCAPIM, C. A.; DO AMARAL JUNIOR, A. T.; DE ARAÚJO RODOVALHO, M.; VIEIRA, R. A.; SCHUELTER, A. R. Avaliação da capacidade de combinação de famílias S2 de milho-pipoca por meio de diferentes testadores. **Semina: Ciências Agrárias**, v. 33, n. 3, 2012.
- BATISTA, B. Soluções de Equações Diferenciais Usando Redes Neurais de Múltiplas camadas com os métodos da Descida mais íngreme e Levenberg-Marquardt. **Ppgme.Ufpa.Br**, p. 90, 2012.

- BESPALHOK F., J.C.; GUERRA, E.P.; OLIVEIRA, R. **Métodos dos retrocruzamentos**. In: BORÉM, A; LIMA, R, O. Melhoramento de Milho. Viçosa: Universidade Federal de Viçosa, ed. 3, p. 120-130, 2018.
- BERLAN, J. Hybrid corn and the unsettled question of heterosis. **Journal of genetics**, v. 97, n. 5, p. 1075-1082, 2018
- BERNARDO, R. Relationship between single-cross performance and molecular marker heterozygosity. **Theoretical and Applied Genetics**, Berlin, v. 83, n.5, p.628-634, 1992.
- BERNINI, C. S.; PATERNIANI, M. E. A. G. Z. Estimativas de parâmetros de heterose em híbridos de populações F2 de milho. **Pesquisa Agropecuária Tropical**, v. 42, n. 1, p. 56-62, 2012.
- BESPALHOK, F.; GUERRA, E. P.; OLIVEIRA, R. **Melhoramento de populações por meio de seleção**. In: BORÉM A., MIRANDA G.V., FRISCHE-NETO R. Melhoria de plantas. v.7, p. 11-20, 2017
- BEYENE, Y.; GOWDA, M.; SURESH, L. M.; MUGO, S.; OLSEN, M.; OIKEH, S. O.; ... PRASANNA, B. M. Genetic analysis of tropical maize inbred lines for resistance to maize lethal necrosis disease. **Euphytica**, v. 213, n. 9, p. 224, 2017.
- BORÉM, A. **Melhoramento de plantas**. 2ed. Viçosa: UFV, p. 453, 1998.
- BORÉM, A.; CAIXETA, E.T. **Marcadores moleculares**. 2 ed. Viçosa: UFV, p. 374, 2009.
- BORÉM, A.; MIRANDA, G.V. **Melhoramento de plantas**. 6. ed. Viçosa-MG: UFV, p. 523, 2013.
- BRAMMER, S. P. **Marcadores moleculares: princípios básicos e uso em programas de melhoramento genético vegetal**. Passo Fundo: Embrapa Trigo, (Embrapa Trigo. Doc. Online 3). 2000. Disponível em: <http://www.cnpt.embrapa.br/biblio/p_do03.htm>. Acesso em 20 de jun. de 2020.
- BUENO, L. C. S.; MENDES, A. N. G.; CARVALHO, S. P. **Melhoramento genético de plantas: princípios e procedimentos**. Lavras: Editora da Universidade Federal de Lavras, p. 175 -190, 2006.
- BUSANELLO, C. **Avaliação de caracteres quantitativos através da análise dialética e mapeamento de QTLs em milho**. 2016. Tese (Doutorado em Agronomia). Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul.
- CAIXETA, E. T.; FERRÃO, L. F. V.; ZAMBOLIM, M. E. Marcadores moleculares. In: BORÉM, A.; FRITSCHÉ-NETO, R. **Biotecnologia aplicada ao melhoramento de plantas**. Viçosa, MG: Suprema, cap.2, p. 336, 2013.
- CARENA, J. C.; BARI M. A. A. **Germoplasma e Grupos Heteróticos**. In: BORÉM, A; LIMA, R, O. Melhoria de Milho. Viçosa: Universidade Federal de Viçosa, ed. 3. p. 47-53, 2018.

CARLONE, M. R.; RUSSELL, W. A. Evaluation of S2 maize lines reproduced for several generations by random mating within lines: Comparisons between the original and maintained S2 lines. **Crop Science**, Madison, v. 28, n. 6, p. 916-20, 1988

CARVALHO, A. D. F.; SOUZA, J. C.; RIBEIRO, P. H. E. Desempenho de híbridos de linhagens parcialmente endogâmicas de milho em regiões dos Estados de Roraima e Minas Gerais. **Ciência e Agrotecnologia**. Lavras, v. 27, n. 5, out. 2003.

CARVALHO, A. D. F.; SOUZA, J. C.; RAMALHO, M. A. P. Capacidade de combinação de progênies parcialmente endogâmicas obtidas de híbridos comerciais de milho. **Revista Brasileira de Milho e Sorgo**, v. 3, n. 3, 2010.

CARVALHO, A. D. F. Capacidade de combinação de linhagens parcialmente endogâmicas obtidas de híbridos comerciais de milho. 2004. **Tese (Mestrado em Genética e Melhoramento de Plantas) Universidade Federal de Lavras (UFLA), Lavras, São Paulo.**

CECCARELLI, S.; GRANDO, S.; MAATOUGUI, M.; MICHAEL, M.; SLASH, M.; HAGHPARAST, R.; ... LABDI, M. Plant breeding and climate changes. **The Journal of Agricultural Science**, v. 148, n. 6, p. 627-637, 2010.

COMPANHIA NACIONAL DE ABASTECIMENTO – CONAB. **Perspectiva agropecuária**. Brasília, 2019. Disponível em: <<https://www.conab.gov.br/info-agro/safras>>, Acesso em: 2 de ago. 2020.

DERAL. **Boletim diário**: Boletins informativos atuais: Milho: análise da Conjuntura. Curitiba. 2018. Disponível em: <http://www.agricultura.pr.gov.br/modules/conteudo/conteudo.php?conteudo=240>. Acesso em: 8 mai. 2020.

FALCONER, D. S. **Introduction to quantitative genetics**. New York. British Library, 1981.

FALK, K. G.; JUBERY, T. Z.; MINEZAMI, S. V.; PARMLY, K. A.; SARKAR, S.; SINGHT, A.; ... SINGH, A. K. Computer vision and machine learning enabled soybean root phenotyping pipeline. **Plant methods**, v. 16, n. 1, p. 5-18, 2020.

FAO. FAOSTAT: Producción. Roma, 2019. Disponível em: <<http://www.fao.org/faostat/en/#data/QC/visualize>>. Acesso em: 13 mar 2020.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores RAPD e RFLP em análise genética**. Brasília: EMBRAPA-CENARGEN, ed. 3, p. 220, 1998.

FORMIGONI, I. **Estimativa de produção de milho na safra 2018/19, por país**. Disponível em: <<http://www.farmnews.com.br/mercado/estimativa-de-producao-de-milho-3/>>. Acesso em: 20/07/2020

FORNASIERI FILHO, D. **Manual da cultura do milho**. Jaboticabal: Funep, 2007.

FREDERIZZI, L. C. Estrutura de um programa de melhoramento de plantas e possíveis aplicações de marcadores moleculares: visão do melhorista. In: **Marcadores moleculares em plantas**. Porto Alegre, p. 2-15, 1998

- GARBUGLIO, D.; DE ARAÚJO, P. M. Potencial genético de linhagens parcialmente endogâmicas e híbridos comerciais. **Journal of Health Sciences**, Londrina, v. 6, n. 1, p. 77-83, 2004
- GARBUGLIO, D. D.; DE ARAÚJO, P. M.. Potencial genético de linhagens parcialmente endogâmicas e híbridos comerciais. **Journal of Health Sciences**, 2015.
- GARCÍA-MARTÍNEZ, H.; FLORES-MAGDALENO, H.; ASCENCIO-HERNÁNDEZ, R., KHALI-GARDEZI, A.; TIJERINA-CHÁVEZ, L.; MANCILLA-VILLA, O. R.; VÁZQUEZ-PENÃ, M. A. Corn Grain Yield Estimation from Vegetation Indices, Canopy Cover, Plant Density, and a Neural Network Using Multispectral and RGB Images Acquired with Unmanned Aerial Vehicles. **Agriculture**, v. 10, n. 7, p. 277, 2020.
- GAYOSSO-BARRAGÁN, O.; RODRÍGUEZ-HERRERA, S. A.; PETROLI, C. D.; ANTUNA-GRIJALVA, O.; LÓPEZ-BENÍTEZ, A.; MANCERA-RICO, A., ... LOZANO-DEL RÍO, A. J. Genetic components for fodder yield and agronomic characters in maize lines. **Agronomy Research**, v. 18, n. 1, p. 77-87, 2020.
- GIORDANI, W.; SCAPIM, C. A.; RUAS, P. M.; RUAS, C. D. F.; CONTRERAS-SOTO, R., Coan, M., ... Gonçalves, L. S. A. Genetic diversity, population structure and AFLP markers associated with maize reaction to southern rust. **Bragantia**, v. 78, n. 2, 2019.
- GRASSI, P.G. M.; DA COSTA ANDRADE, J. A.; SILVA, C. L. S. P.; DE MORAES, M. L. T.; CERESINI, P. C. Divergência genética entre linhagens de milho utilizando microssatélite e correlação com desempenho de híbrido simples. **Cultura Agrônômica: Revista de Ciências Agrônômicas**, v. 24, n. 1, p. 1-16, 2015.
- GRATTAPAGLIA, D. Marcadores moleculares em espécies florestais: Eucalyptus como modelo. In: NASS, L.L.; VALOIS, A.C.C.; MELO, I.S.; VALADARES-ILGLIS, M.C. **Recursos genéticos e melhoramento: plantas**. Rondonópolis: Fundação MT, p.967- 1010, 2001.
- GONÇALVES, M. C.; DAVIDE, L. M. C. Desempenho produtivo de linhagens parcialmente endogâmicas S₂. **Revista de Ciências Exatas e da Terra UNIGRAN**, v. 2, n. 1, 2013.
- GONZÁLEZ-CAMACHO, J. M.; ORNELLA, L.; PÉREZ-RODRÍGUEZ, P.; GIANOLA, D., DREISIGACKER, S.; CROSSA, J. Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. **The plant genome**, v. 11, n. 2, p. 1-15, 2018.
- GOULET, B. E.; RODA, F.; HOPKINS, R. Hybridization in plants: old ideas, new techniques. **Plant physiology**, v. 173, n. 1, p. 65-78, 2017.
- GUIMARÃES, C. T. et al. **Mapeamento de QTLs e Seleção Assistida por Marcadores Moleculares**. In: BORÉM, A; LIMA, R, O. Melhoramento de Milho. Viçosa: Universidade Federal de Viçosa, ed. 3. p. 307-320, 2018.
- HALLAUER, A. R.; CARENA, M. J.; MIRANDA FILHO, J B de. Quantitative genetics in maize breeding. **Springer Science & Business Media**, v. 6, 2010.

HALLAUER, A.R.; MIRANDA FILHO, J. B. **Quantitative genetics in maize breeding**. Ames: Iowa State University Press, p. 488, 1995

JHA, K.; DOSHI, A.; PATEL, P.; SHAH, M. A comprehensive review on automation in agriculture using artificial intelligence. **Artificial Intelligence in Agriculture**, v. 2, p. 1-12, 2019.

JIANG, S.; CHENG, Q.; YAN, J.; FU, R.; WANG, X. Genome optimization for improvement of maize breeding. **Theoretical and Applied Genetics**, v. 133, p. 1491–1502, 2020.

JUNIOR, O. P.; FARIA, M. V.; BATTISTELLI, G. M.; ROSSI, E. S.; DE MARCK, D. F.; DA SILVA, C. A. Avaliação de linhagens S₂ de milho em topcrosses com linhagem-elite testadora. **Revista Brasileira de Milho e Sorgo**, v. 16, n. 2, p. 297-309, 2017.

KAGE, U.; MADALAGERI, D.; MALAKANNAVAR, L.; GANAGASHETTY, P. Genetic diversity studies in newly derived inbred lines of maize (*Zea mays* L.). **Molecular Plant Breeding**, v. 4, 2013.

KAPOOR, C.; LATA, S.; SHARMA, J. K. Combining ability and heterosis studies for grain yield and its component traits in maize (*Zea mays* L.). **Electronic Journal of Plant Breeding**, v. 5, n. 4, p. 716-721, 2014.

LANA, M. A.; EULENSTEIN, F.; SCHLINDWEIN, S. L.; GRAEF, F.; SIEBER, S.; BITTENCOUR, H. V. H. Yield stability and lower susceptibility to abiotic stresses of improved open-pollinated and hybrid maize cultivars. **Agronomy for Sustainable Development**, v. 37, n. 4, p. 30-41, 2017.

LEGESSE, B. W.; MYBURG, A. A.; PIXLEY, K. V.; TWUMASI-AFRIYIE, S., BOTHA, A. M. Relationship between hybrid performance and AFLP based genetic distance in highland maize inbred lines. **Euphytica**, v. 162, n. 3, p. 313-323, 2008.

Ministério da Agricultura, Pecuária e Abastecimento – MAPA. **Projeções do agronegócio**. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/politica-agricola/todas-publicacoes-de-politica-agricola/projecoes-do-agronegocio/projecoes-do-agronegocio-2018-2019-2028-2029>. Acesso em: 5 mai. 2020.

MAKUMBI, D.; BETRÁN, J. F.; BÄNZIGER, M.; RIBAUT, J. M. Combining ability, heterosis and genetic diversity in tropical maize (*Zea mays* L.) under stress and non-stress conditions. **Euphytica**, v. 180, n. 2, p. 143-162, 2011.

MEIRELLES, P. G. **Divergência genética entre linhagens de milho estimada por microssatélites e correlação com desempenho de híbridos simples**. 2009. Tese (Doutorado em Agronomia), Universidade Estadual Paulista (UNESP), São Paulo.

MENDOZA, A.; NOLDIN, O. Habilidad combinatoria de híbridos de maíz en cruzamiento dialelo. **Tecnologia Agrária**, v. 1, n. 1, p. 32-40, 2016.

MEENA, A. K., GURJAR, D., PATIL, S. S., KUMHAR, B. L. Concept of heterotic group and its exploitation in hybrid breeding. *Int J Curr Microbiol App Sci*, v. 6, p. 61-73, 2017.

- MIURA, J. C.; BERRES, V. A.; TREVIZAN, C. B.; SHIROTA, L. Y.; DAVIDE, L. M. C.; CANDIDO, L. S. **Análise de divergência genética em linhagens parcialmente endogâmicas de milho em Dourados-MS**. Encontro de Ensino Pesquisa e Extensão, Dourados, 2014.
- MOLIN, D.; COELHO, C. J.; MÁXIMO, D. S.; FERREIRA, F. S.; GARDINGO, J. R.; MATIELLO, R. R. Genetic diversity in the germplasm of tropical maize landraces determined using molecular markers. **Genetics and Molecular Research**, v. 12, n. 1, p. 99-114, 2013.
- MOSAFA, S., HUSSEIN, B. A., SAYED, H. A., ELLTRIBY, H. A., HUSSEIN, E. H. Genetic diversity assessment among some ficus species using morphological characters and AFLP. **Plant Archives**. v 20, n. 1, p. 1395-1404, 2020.
- MÔRO, G. V., **Histórico de Melhoramento Genético do Milho**. In: BORÉM, A; LIMA, R, O. Melhoramento de Milho. Viçosa: Universidade Federal de Viçosa, ed. 3, p. 9-20, 2018.
- MURTADHA, M. A.; ARIYO, O. J.; ALGHAMDI, S. S. Analysis of combining ability over environments in diallel crosses of maize (*Zea mays*). **Journal of the Saudi Society of Agricultural Sciences**, v. 17, n. 1, p. 69-78, 2018.
- NADEEN, M. A.; NAWAZ, M. A.; SHAHID, M. Q.; DOGAN, Y.; COMERTPAY, G.; YILDIZ, M., ...OZKAN, H. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. **Biotechnology & Biotechnological Equipment**, v. 32, n. 2, p. 261-285, 2018.
- NASS, L. L.; NISHIKAWA, M. A. N. Pré-melhoramento de germoplasma vegetal. In: GOES, M (Org.) **Tópicos em recursos genéticos**. Brasília: Embrapa Recursos Genéticos e Biotecnologia, 2001.
- OLIVEIRA de, D. T. **Um sistema inteligente que prevê a produtividade do algodão em imagens de lavouras comerciais**. 2019. Tese (Doutorado em Agronomia), Universidade Estadual Paulista (UNESP), São Paulo.
- OSEI, M. K.; PREMPEH, R.; ADJEBENG-DANQUAH, J.; OPOKU, J. A.; DANQUAH, A.; DANQUAH, E.; ... ADU-DAPAAH, H. Marker-Assisted Selection (MAS): A Fast-Track Tool in Tomato Breeding. In: **Recent Advances in Tomato Breeding and Production**. IntechOpen, 2018.
- PALAPARTHI, D. ; WALI, M. C. Genetic variability and heritability for yield and its traits in diverse S2 populations developed through reciprocal recurrent selection in maize (*Zea mays* L.). **International Journal of Chemical Studies: IJCS**, v. 7, n. 3, p. 2668-2671, 2019.
- PARMLEY, K. A.; HIGGINS, R. H.; GANAPATHYSUBRAMANIAN, B.; SARKAR, S.; SINGH, A. K. Machine learning approach for prescriptive plant breeding. **Scientific reports**, v. 9, n. 1, p. 1-12, 2019.
- PATERNIANI, M. E. A. G. Z.; BERNINI, C. S. Estimativas de parâmetros de heterose em híbridos de populações F₂ de milho. **Pesquisa Agropecuária Tropical**. Goiânia, v. 42, n. 1, 2012.

- PATERNIANI, M. E. A. G. Z.; GUIMARÃES, P. D. S.; LURDES, R. R.; GALLO, P. B.; SOUZA, A. P. D.; LABORDA, P. R.; OLIVEIRA, K. M. Capacidade combinatória, divergência genética entre linhagens de milho e correlação com heterose. **Bragantia**, v. 67, n. 3, 2008.
- PATERNIANI, E.; MIRANDA FILHO, J. B. de. Melhoramento de populações. **Melhoramento e produção de milho**. Campinas: Fundação Cargill, v. 1, p. 217-264, 1987.
- PRADO, S. A.; GAMBIN, B. L.; NOVOA, A. D.; FOSTER, D.; SENIOR, M. L.; ZINSELMEIER, C.; ... BORRÁS, L. Correlations between parental inbred lines and derived hybrid performance for grain filling traits in maize. **Crop Science**, v. 53, n. 4, p. 1636-1645, 2013.
- QI, X.; KIMATU, J. N.; LI, Z.; JIANG, L.; CUI, Y.; LIU, B. Heterotic analysis using AFLP markers reveals moderate correlations between specific combining ability and genetic distance in maize inbred lines. **African Journal of Biotechnology**, v. 9, n. 11, p. 1568-1572, 2010.
- QIN, Z.; MYRERS, D. B.; RANSOM, C. J.; KITCHEN, N. R.; LIANG, S. Z.; CAMBERATO, J. J.; ... SHANAHAN, J. F. Application of machine learning methodologies for predicting corn economic optimal nitrogen rate. **Agronomy Journal**, v. 110, n. 6, p. 2596–2607, 2018.
- RAMÍREZ, M. E.; CARBALLO, A. C.; AGUILAR, C. H.. Endogamic Advancement of S₂ maize lines based on differential responses to magnetic field treatment. **African Journal of Agricultural Science and Technology (AJAST)**, v. 4, n. 3, p. 624-632, 2016.
- RIGON, J. P. G.; CAPUANI, S.; RIGON, C. A.G. Genetic divergence among maize hybrids by morphological descriptors. **Bragantia**, v. 74, n. 2, p. 156-160, 2015.
- RITTE, I.; EGNIN, M.; KUSOLWA, P.; BERNARD, G. C.; HE, G.; BONSI, C. Characterization of markers linked to resistance motifs against maize lethal necrosis in Tanzanian maize germplasm. **African Journal of Plant Science**, v. 12, n. 10, p. 238-248, 2018.
- ROY, N. S.; KIM, N.S. Genetic diversity analysis of maize lines using AFLP and TE-based molecular marker systems. **Genes & Genomics**, v. 38, n. 10, p. 1005-1012, 2016.
- SHANMUGANATHAN, S. Artificial neural network modelling: An introduction. In: Artificial neural network modelling. **Springer**, v. 628, n. 1, p. 1-14, 2016.
- SHARMA, S.; SHARMA, A. Molecular markers based plant breeding. **Advances in Research**, v. 16, n.1, p. 1-15, 2018.
- SJARIF, N. N. A., AZMI, N. F., SARKAN, H. M., SAM, S. M., OSMAN, M. Z. Predicting Churn: How Multilayer Perceptron Method Can Help with Customer Retention in Telecom Industry. In **IOP Conference Series: Materials Science and Engineering**, v. 864, n 1, 2020
- SILVA, G. N.; TOMAZ, R. S.; SANT'ANNA, I. C.; CARNEIRO, V. Q.; CRUZ, C. D.; NASCIMENTO, M. Evaluation of the efficiency of artificial neural networks for genetic value prediction. **Genetic Molecular Research**, v. 15, p. 1-11, 2016.

- SILVA, C. A.; FARIA, M. V.; ROSSI, E. S.; JUNIOR, O. P.; BATTISTELLI, G. M.; GABRIEL, A.; DE MARCK, D. F. Performance and stability of maize topcross hybrids from partly inbred lines. **Semina: Ciências Agrárias**, v. 38, n. 5, p. 2899-2908, 2017.
- SILVA, R. M.; MIRANDA FILHO, J. B. Heterosis expression in crosses between maize populations: ear yield. **Scientia Agricola**, v. 60, n. 3, p. 519-524, 2003.
- SINGH, P.; SALGOTRA, S. K.; SINGH, A. K.; SHARMA, M.; GUPTA, A.; SALGOTRA, S. K.; ... SHARMA, M. Stability and genetic divergence study of single cross hybrids in maize (*Zea mays* L.). **African Journal of Agricultural Research**, v. 10, n. 31, p. 3080-3085, 2015.
- SINIMBU, F. **A importância do melhoramento genético na Cultura do milho**, 2015. Disponível em: <<http://www.paginarural.com.br/artigo/705/aimportanciadomelhoramentogeneticonacultu radomilho>>. Acesso em: 4 fev. 2020.
- SOUZA, J. C **Cultivares**. In: BORÉM, A; LIMA, R, O. Melhoramento de Milho. Viçosa: Universidade Federal de Viçosa, ed. 3. 2018. p. 295-302.
- SOUZA, G.M.; BARBOSA, A.M. **Fatores de estresse no milho são diversos e exigem monitoramento constante**. ESALQ, Piracicaba, 2013.
- SOUZA JÚNIOR, C. L. **Melhoramento de espécies alógamas**. In: NASS, L. L. Recursos genéticos e melhoramento de plantas. Rondonópolis: Fundação MT, p. 159-200, 2001.
- STORM, H.; HECKELEI, T.; BAYLIS, K.; MITTENZWEI, K. Identifying effects of farm subsidies on structural change using neural networks. **Agricultural and Resources Economics, Discusison Paper**. Bonn: Institute for Food and Resource Economics, 2019. Disponível em: http://www.ilr.uni-bonn.de/agpo/publ/dispap/download/dispap19_01.pdf. Acesso em: 20 julho 2020.
- SU, C.; WANG, W.; GONG, S.; ZUO, J.; LI, S.; XU, S. High density linkage map construction and mapping of yield trait QTLs in maize (*Zea mays*) using the genotyping-by-sequencing (GBS) technology. **Frontiers in plant science**, v. 8, p. 706, 2017.
- TSEHAYE, H.; ELAMEEN, A.; TRONSMO, A. M.; SUNDHEIM, L.; TRONSMO, A.; ASSEFA, D.; BRURBERG, M. B. Genetic variation among *Fusarium verticillioides* isolates associated with Ethiopian maize kernels as revealed by AFLP analysis. **European journal of plant pathology**, v. 146, n. 4, p. 807-816, 2016.
- UECE. **Inteligência artificial**. Fortaleza: Computação/Informática, Universidade Estadual do Ceará, 2017. Disponível em: http://www.uece.br/computacaoead/index.php/downloads/doc_view/2177-inteligencia-artificial?tmpl=component&format=raw. Acesso em: 25 maio 2020.
- USDA - United State Department of Agriculture. **Agriculture Research Service**. Washington, 2019. Disponível em: <<https://www.ars.usda.gov/oc/timeline/corn/>> Acesso em: 15 de jul. 2019.

VENCOVSKY, R. Herança quantitativa. **Melhoramento e produção de milho**, v. 2, p. 137-214, 1987.

VETTORAZI, J. C. F. **Seleção recorrente recíproca em milho (*Zea Mays L.*) monitorada por marcadores SSR–EST's associados a produtividade**. 2016. Tese de Doutorado. Dissertação (Mestrado em Genética e Melhoramento de Plantas)–Campos dos Goytacazes–RJ, Universidade Estadual do Norte Fluminense-UENF, 90p.

VIAN, A. L.; SANTI, A. L.; AMADO, T. J. C.; CHERUBIN, M. R.; SIMON, D. H., DAMIAN, J. M.; BREDEMEIER, C. Variabilidade espacial da produtividade de milho irrigado e sua correlação com variáveis explicativas de planta. **Ciência Rural**, v. 46, n. 3, p. 464-471, 2016.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; LEE, T. V. D.; HORNES, M.; ... ZABEAU, M. AFLP: a new technique for DNA fingerprinting. **Nucleic acids research**, v. 23, n. 21, p. 4407-4414, 1995.

VUYLSTEKE, M.; MANK, R.; ANTONIESE, R.; BASTIAANS, E.; SENIOR, M. L.; STUBER, C. W.; ... ZABEAU, M. Two high-density AFLP® linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. **Theoretical and Applied Genetics**, v. 99, p. 921–935, 1999

WANG, S. H., ZHANG, Y., LI, Y. J., JIA, W. J., LIU, F. Y., YANG, M. M., ZHANG, Y. D. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. **Multimedia Tools and Applications**, v. 77, n. 9, p. 10393-10417, 2018.

WU, W. D., LIU, W. H., SUN, M., ZHOU, J. Q., LIU, W., ZHANG, C. L., ... & MA, X. Genetic diversity and structure of *Elymus tangutorum* accessions from western China as unraveled by AFLP markers. **Hereditas**, v. 156, n. 1, p. 8, 2019.

ZYKOVA, M. V.; BRAZOVSKY, K. S.; VERETENNIKOVA, E. E.; DANILETS, M. G.; LOGVINOVA, L. A.; ROMANENKO, S. V.; ... LYAPKOY, A. A. New artificial network model to estimate biological activity of peat humic acids. **Environmental Research**, v. 187, n. 1, p. 1-15, 2020.

XIA, X. C.; REIF, J. C.; HOISINGTON, D. A.; MELCHINGER, A. E.; FRISCH, M.; WARBURTON, M. L. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers. **Crop Science**, v. 45, n. 6, p. 2573-2582, 2005.

Referências do artigo A

BESPALHOK, F.; GUERRA, E. P.; OLIVEIRA, R. **Melhoramento de populações por meio de seleção.** In: BORÉM A., MIRANDA G.V., FRISCHE-NETO R. Melhoramento de plantas. v.7, p. 11-20, 2017.

BEYENE, Yoseph et al. Genetic analysis of tropical maize inbred lines for resistance to maize lethal necrosis disease. **Euphytica**, v. 213, n. 9, p. 224, 2017.

CAIXETA, E. T.; FERRÃO, L. F. V.; ZAMBOLIM, M. E. **Marcadores moleculares.** In: BORÉM, A.; FRITSCHÉ-NETO, R. Biotecnologia aplicada ao melhoramento de plantas. Viçosa, MG: Suprema, cap.2, p. 336, 2013.

CARLONE, M. R.; RUSSELL, W. A. Evaluation of S2 maize lines reproduced for several generations by random mating within lines: Comparisons between the original and maintained S2 lines. **Crop Science**, v. 28, n. 6, p. 916-20, 1988

CORANDER, J., WALDMANN, P., MARTTINEN, P. and SILLANPÄÄ, M.J. BAPS 2: enhanced possibilities for the analysis of genetic population structure, **Bioinformatics**, v. 20, p. 2363-2369, 2004.

COSTA, R., PEREIRA, G., GARRIDO, I., SOUZA, M. M.T., ESPINOSA, F. Comparison of RAPD, ISSR, and AFLP molecular markers to reveal and classify orchardgrass (*Dactylis glomerata* L.) germplasm variations. **PloS one**,/ v. 11, n. 4, 2016.

DOYLE, J. J.; DOYLE, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochemical Bulletin**, v. 19, p. 11-15, 1987.

EXCOFFIER, L., LAVAL, G., SCHNEIDER, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. **Evolutionary bioinformatics**, v. 1, p. 47-50, 2005.

FERREIRA, E. B.; CAVALCANTI, P. P.; NOGUEIRA, D. A. ExpDes. pt: Experimental Designs package. R package version 1.2. 0. **Retrieved Sept**, v. 11, p. 2018, 2018.

FRASCAROLI, E., SCHRAG, T. A., MELCHINGER, A. E. Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. **Theoretical and applied genetics**, v. 126, n. 1, p. 133-141, 2013.

GADHERI, A.; ADAMS, M.W.; NASSIB, A.M. Relationship between genetic distance and heterosis for yield and morphological traits in dry Edible Bean and Faba Bean. **Crop Science**, v.24, p.37-42, 1984.

GARBUGLIO, D. D., GERAGE, A. C., ARAÚJO, P. M. D., FONSECA JUNIOR, N. D. S., SHIOGA, P. S. Análise de fatores e regressão bissegmentada em estudos de estratificação ambiental e adaptabilidade em milho. **Pesquisa Agropecuária Brasileira**, v. 42, n. 2, p. 183-191, 2007.

- GIORDANI, W.; SCAPIM, C. A.; RUAS, P. M.; RUAS, C. D. F.; CONTRERAS-SOTO, R., Coan, M., ... Gonçalves, L. S. A. Genetic diversity, population structure and AFLP markers associated with maize reaction to southern rust. **Bragantia**, v. 78, n. 2, 2019.
- GUIMARÃES, P. D.; PATERNIANI, M. E. A. G. Z.; LUDERS, R. R.; DE SOUZA, A. P.; LABORDA, P. R.; OLIVEIRA, K. M. Correlation between the heterosis of maize hybrids and genetic divergence among lines. **Pesquisa Agropecuária Brasileira**, v. 23, n. 6, 2007
- HUSON D. H. and BRYANT D. Application of Phylogenetic Networks in Evolutionary Studies, **Mol. Biol. Evol.** v. 23, n. 2, p. 254-267, 2006.
- JUNIOR, O. P., FARIA, M. V., BATTISTELLI, G. M., ROSSI, E. S., DE MARCK, D. F., & DA SILVA, C. A. Avaliação de linhagens S2 de milho em topcrosses com linhagem-elite testadora. **Revista Brasileira de Milho e Sorgo**, v. 16, n. 2, p. 297-309, 2017.
- KRONKA, S. N.; BANZATO, D. A. Estatística: sistema para análise estatística versão 2.0. Jaboticabal, FUNEP. p. 247, 1995.
- MIRANDA FILHO Filho, J. B., VENCOSKY, R. The partial circulant diallel cross at the interpopulation level. **Genetics and molecular biology**, v. 22, n. 2, p. 249-255, 1999.
- MOLIN, D. et al. Genetic diversity in the germplasm of tropical maize landraces determined using molecular markers. **Genet. Mol. Res**, v. 12, n. 1, p. 99-114, 2013.
- MOSAFSA, S., HUSSEIN, B. A., SAYED, H. A., ELLTRIBY, H. A., HUSSEIN, E. H. Genetic diversity assessment among some ficus species using morphological characters and AFLP. **Plant Archives**. v 20, n. 1, p. 1395-1404, 2020.
- OLIVOTO, T.; NARDINO, M.; CARVALHO, I. R.; FOLLMANN, D. N.; FERRARI, M.; SZARESKI, V. J., ... DE SOUZA, V. Q. REML/BLUP and sequential path analysis in estimating genotypic values and interrelationships among simple maize grain yield-related traits. **Genetics and Molecular Research**, v. 16, n. 1, p. 1-19, 2017.
- PELEGRIN, A. J.; CARVALHO, I. R.; NUNES, A. C. P.; DEMARI, G. H.; SZARESKI, V. J.; BARBOSA, M. H., ... DE RESENDE, M. D. V. Adaptability, stability and multivariate selection by mixed models. **American Journal of Plant Sciences**, v. 8, n. 13, p. 3324, 2017.
- PIMENTEL-GOMES, F. Course of experimental statistics. **Piracicaba: FEALQ**, v. 15, 2000.
- R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. R Found. **Stat. Comput. Vienna, Austria**, 2017
- RAMÍREZ, M. E.; CARBALLO, A. C.; AGUILAR, C. H. Endogamic advancement of S2 maize lines based on differential responses to magnetic field treatment. **African Journal of Agricultural Science and Technology (AJAST)**, v. 4, n. 3, p. 624-632, 2016.
- RIGON, J. P. G.; CAPUANI, S.; RIGON, C. A.G. Genetic divergence among maize hybrids by morphological descriptors. **Bragantia**, v. 74, n. 2, p. 156-160, 2015.

RODRÍGUEZ-PEÑA, R. A., JOHNSON, R. L., JOHNSON, L. A., ANDERSON, C. D., RICKS, N. J., FARLEY, K. M., ... STEVENS, M. R. Investigating the genetic diversity and differentiation patterns in the *Penstemon scariosus* species complex under different sample sizes using AFLPs and SSRs. **Conservation Genetics**, v.19, n. 6, p. 1335-1348, 2018.

SHARMA, P. V. K.; KUMAR, P. Microsatellite marker dependent genetic divergence assessment within and among heterotic groups of tropical maize inbred lines. **Maydica**, v. 63, n. 3, p. 13, 2018.

STONE, B. W., Ward, A., FARENWALD, M., LUTZ, A. W., WOLFE, A. D. Genetic diversity and population structure in Cary's Beardtongue *Penstemon caryi* (Plantaginaceae), a rare plant endemic to the eastern Rocky Mountains of Wyoming and Montana. **Conservation Genetics**, v. 20, n. 5, p. 1149-1161, 2019.

VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M., LEE, T. VAN DE, HORNES, M., Friters, A., POT, J., PALEMAN, J. and KUIPER, M. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, v. 23, p. 4407-4414, 1995.

WU, W. D., LIU, W. H., SUN, M., ZHOU, J. Q., LIU, W., ZHANG, C. L., ... & MA, X. Genetic diversity and structure of *Elymus tangutorum* accessions from western China as unraveled by AFLP markers. **Hereditas**, v. 156, n. 1, p. 8, 2019.

Referências - Artigo b

- AZEVEDO, A. M., ANDRADE JUNIOR, V. C. D., PEDROSA, C. E., OLIVEIRA, C. M. D., DORNAS, M. F. S., CRUZ, C. D., VALADARES, N. R. Application of artificial neural networks in indirect selection: a case study on the breeding of lettuce. **Bragantia**, v. 74, p. 387-394, 2015.
- AZODI, C. B., BOLGER, E., MCCARREN, A., ROANTREE, M., DE LOS CAMPOS, G., SHIU, S. H. Benchmarking parametric and Machine Learning models for genomic prediction of complex traits. **G3: Genes, Genomes, Genetics**, v. 9, n. 11, p. 3691-3702, 2019.
- BALESTRE, M., VON PINHO, R. G., SOUZA, J. C., LIMA, J. L. Comparison of maize similarity and dissimilarity genetic coefficients based on microsatellites markers. **Genetics and Molecular Research**, v. 7, n. 3, p. 695-705, 2008.
- BERGAMASCHI, H., DALMAGO, G. A., BERGONCI, J. I., BIANCHI, C. A. M., MULLER, A. G., COMIRAN, F., HECKLER, B. M. M. Distribuição hídrica no período crítico do milho e produção de grãos. **Pesquisa Agropecuária Brasileira**, v. 39, n. 9, p. 831-839, 2004.
- BESPALHOK, F.; GUERRA, E. P.; OLIVEIRA, R. Melhoramento de populações por meio de seleção. In: BORÉM A., MIRANDA G.V., FRISCHE-NETO R. **Melhoramento de plantas**, v.7, 2017. p. 11-20.
- DEB, M., OGUNFUNMI, T. Using information theoretic learning techniques to train neural networks. In: **51st Asilomar Conference on Signals, Systems, and Computers. IEEE**. 2017. p. 351-355.
- DOYLE, J. J., DOYLE, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochemical Bulletin**, v. 19, p. 11-15, 1987.
- ELI-CHUKWU, N. Applications of artificial intelligence in agriculture: A review. **Engineering, Technology & Applied Science Research**, v. 9, n. 4, p. 4377-4383, 2019.
- FALK, K. G., JUBERY, T. Z., O'ROURKE, J. A., SINGH, A., SARKAR, S., GANAPATHYSUBRAMANIAN, B., SINGH, A. K. Soybean Root System Architecture Trait Study through Genotypic, Phenotypic, and Shape-Based Clusters. **Plant Phenomics**, v. 20, p. 1-23, 2020.
- FANCELLI, A.L.; DOURADO-NETO, D. **Produção de milho**. Guaíba: Livraria e Editora Agropecuária, 2000. 360 p.
- GEDIL, M., MENKIR, A. An integrated molecular and conventional breeding scheme for enhancing genetic gain in maize in Africa. **Frontiers in Plant Science**, v. 10, p. 1-17, 2019.
- GONZÁLEZ-CAMACHO, J. M., ORNELLA, L., PÉREZ-RODRÍGUEZ, P., GIANOLA, D., DREISIGACKER, S., CROSSA, J. Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. **The plant genome**, v. 11, n. 2, p. 1-15, 2018.

- HESLOT, N., YANG, H. P., SORRELLS, M. E., JANNINK, J. L. Genomic selection in plant breeding: a comparison of models. **Crop science**, v. 52, n.1, p. 146-160, 2012.
- JHA, K.; DOSHI, A.; PATEL, P.; SHAH, M. A comprehensive review on automation in agriculture using artificial intelligence. **Artificial Intelligence in Agriculture**, v. 2, p. 1-12, 2019.
- JIANG, S., CHENG, Q., YAN, J., FU, R., WANG, X. Genome optimization for improvement of maize breeding. **Theoretical and Applied Genetics**, v. 133, p. 1491–1502, 2020.
- KAAB, A., SHARIFI, M., MOBIL, H., NABAVI-PELESARAEI, A., CHAU, K. W. Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. **Science of the Total Environment**, v. 664, p. 1005-1019, 2019.
- LANA, M. A., EULENSTEIN, F., SCHLINDWEIN, S. L., GRAEF, F., SIEBER, S., BITTENCOUR, H. V. H. Yield stability and lower susceptibility to abiotic stresses of improved open-pollinated and hybrid maize cultivars. **Agronomy for Sustainable Development**, v. 37, n. 4, p. 30-41, 2017.
- LEDELL, E.; POIRIER, S. H2o automl: Scalable automatic machine learning. In: **Proceedings of the AutoML Workshop at ICML**, 2020.
- MEENA, A. K., GURJAR, D., PATIL, S. S., KUMHAR, B. L. Concept of heterotic group and its exploitation in hybrid breeding. **Int J Curr Microbiol App Sci**, v. 6, p. 61-73, 2017.
- NEHE, A., AKIN, B., SANAL, T., EVLICE, A. K., ÜNSAL, R., DINCER, N., ... YAKTUBAY, S. Genotype x environment interaction and genetic gain for grain yield and grain quality traits in Turkish spring wheat released between 1964 and 2010. **PloS one**, v. 14, n. 7, 2019.
- NIELSEN, R. L. Kernel set scuttlebutt. **Corn News Network**, West Lafayette, IN. 2005. Disponível em: <<http://www.kingcorn.org/news/articles.05/KernelSet-0809.html>>. Acesso em: jun. 2019
- OLIVEIRA, L. F. Equipe Multidisciplinar para a utilização de Inteligência Artificial Aplicada as Imagens Médicas. **Journal of Health Informatics**, v. 11, n. 3, 2019.
- PARMLEY, K. A., HIGGINS, R. H., GANAPATHYSUBRAMANIAN, B., SARKAR, S., SINGH, A. K. Machine learning approach for prescriptive plant breeding. **Scientific reports**, v. 9, n. 1, p. 1-12, 2019.
- PATRÍCIO, D. I., RIEDER, R. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. **Computers and electronics in agriculture**, v. 153, p. 69-81, 2018.
- PIVOTO, D., WAQUIL, P. D., TALAMINI, E., FINOCCHIO, C. P. S., DALLA CORTE, V. F., DE VARGAS MORES, G. Scientific development of smart farming technologies and their application in Brazil. **Information Processing in Agriculture**, V. 5, n. 1, p. 21-32, 2017.

QIAO, P., LIN, M., VASQUEZ, M., MATSCHI, S., CHAMNESS, J., BASEGGIO, M.,.....SCANLON, M. J. Machine Learning Enables High-Throughput Phenotyping for Analyses of the Genetic Architecture of Bulliform Cell Patterning in Maize. **G3: Genes, Genomes, Genetics**, v. 9, n. 12, p. 4235-4243, 2019.

R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. R Found. **Stat. Comput. Vienna, Austria**, 2017.

RAHMANIFARD, H., PLAKSINA, T. Application of artificial intelligence techniques in the petroleum industry: a review. **Artificial Intelligence Review**, v. 52, n. 4, p. 2295-2318, 2019.

RITCHIE, S. W.; HANWAY, J. J.; BENSON, G. O. **How a corn plant develops**. Special Bulletin, Iowa, n. 48. 1993.

SHANMUGANATHAN, S. Artificial neural network modelling: An introduction. In: Artificial neural network modelling. **Springer**, v. 628, p. 1-14. 2016.

SHANMUGANATHAN, S. Artificial neural network modelling: An introduction. In: Artificial neural network modelling. **Springer**, v. 628, n. 1, p. 1-14, 2016.

SHIN, S. Designing the Framework of Evaluation on Learner's Cognitive Skill for Artificial Intelligence Education through Computational Thinking. **Journal of The Korean Association of Information Education**, v. 24, n. 1, p. 59-69, 2020

SILVA, G. N., TOMAZ, R. S., SANT'ANNA, I. C., CARNEIRO, V. Q., CRUZ, C. D., NASCIMENTO, M. Evaluation of the efficiency of artificial neural networks for genetic value prediction. **Genetic Molecular Research**, v. 15, p. 1-11, 2016.

WANI, M. A., KHAN, G. H., GAZAL, A., LONE, R. A. Genetic Purity Analysis in Maize under Temperate Conditions. **Int. J. Curr. Microbiol. App. Sci**, v. 6, n. 9, p. 2710-2722, 2017.