

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
IMUNOLOGIA E DIP/ GENÉTICA E BIOTECNOLOGIA

Flávia Rangel de Souza

**ANÁLISE DA EXPRESSÃO GÊNICA E DESENVOLVIMENTO DE
MARCADORES MOLECULARES NOS TRANSCRIPTOMAS DE
Cenchrus purpureus E *Urochloa humidicola***

Tese de Doutorado

Juiz de Fora

2021

FLÁVIA RANGEL DE SOUZA

**ANÁLISE DA EXPRESSÃO GÊNICA E DESENVOLVIMENTO DE
MARCADORES MOLECULARES NOS TRANSCRIPTOMAS DE
Cenchrus purpureus E *Urochloa humidicola***

Tese de Doutorado do Curso de Pós-Graduação em Ciências Biológicas: Área: Genética e Biotecnologia, para obtenção do Título de Doutor em Ciências Biológicas: Área: Genética e Biotecnologia.

Orientador: Marco Antônio Machado

Co-orientadora: Ana Luisa Sousa Azevedo

Juiz de Fora

2021

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Souza, Flávia Rangel de .

Análise da expressão gênica e desenvolvimento de marcadores moleculares nos transcriptomas de *Cenchrus purpureus* e *Urochloa humidicola* / Flávia Rangel de Souza. -- 2021.

146 f.

Orientador: Marco Antonio Machado

Coorientadora: Ana Luisa Sousa Azevedo

Tese (doutorado) - Universidade Federal de Juiz de Fora, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Ciências Biológicas: Imunologia e Genética, 2021.

1. RNA-seq. 2. Transcriptoma. 3. Genes diferencialmente expressos. 4. Lignina. 5. Microssatélites. I. Machado, Marco Antonio, orient. II. Azevedo, Ana Luisa Sousa, coorient. III. Título.

FLÁVIA RANGEL DE SOUZA

**ANÁLISE DA EXPRESSÃO GÊNICA E DESENVOLVIMENTO DE
MARCADORES MOLECULARES NOS TRANSCRIPTOMAS DE
Cenchrus purpureus E *Urochloa humidicola***

Tese de Doutorado do Curso de Pós-Graduação em Ciências Biológicas, para obtenção do Título de Doutor em Ciências Biológicas na área de Genética e Biotecnologia

Data de aprovação: 15/09/2021

BANCA EXAMINADORA



Prof. Dr. Marco Antônio Machado - Orientador

Instituição: Embrapa Gado de Leite



Dra. Ana Luisa Sousa Azevedo - Co-Orientadora

Instituição: Embrapa Gado de Leite



Prof. Dr. Lyderson Facio Viccini

Instituição: Universidade Federal de Juiz de Fora

Dr. Juarez Campolina Machado

Instituição: Embrapa Gado de Leite



Prof. Dr. Elisson Antonio da Costa Romanel

Instituição: Universidade de São Paulo – Campus Lorena

AGRADECIMENTOS

Agradeço a Deus, primeiramente, pelas vitórias alcançadas durante os 4 anos de doutorado, mesmo com as dificuldades que a pandemia proporcionou. Agradeço à instituição UFJF, pela oportunidade de aperfeiçoar a minha formação. À Fapemig e CNPq, pelo financiamento do projeto realizado; à Embrapa, pela disponibilidade da infra-estrutura para os estudos; à Capes, pelos financiamentos da bolsa de estudos no Brasil e nos Estados Unidos por meio do Doutorado Sanduíche.

Agradeço ao meu orientador, professor doutor Marco Antonio Machado, pela disponibilidade de orientação e de laboratório para a pesquisa. Agradeço, em especial, à doutora Ana Luisa Sousa Azevedo, co-orientadora da tese, responsável por ensinar, esclarecer dúvidas pertinentes ao trabalho, e por confiar em mim um projeto que envolve bioinformática, algo totalmente novo para mim.

Agradeço a oportunidade que tive em aperfeiçoar os estudos nos Estados Unidos, que só foi possível com o financiamento da Capes e a disponibilidade do Dr. Marcio Resende. Durante o ano que passei lá, tive a oportunidade de estar no Sweet Corn Lab, conhecer pessoas de vários países, além de conhecer amigas brasileiras que levo para vida.

Agradeço aos amigos do Laboratório de Genética Molecular da Embrapa Gado de Leite por me ajudar nessa caminhada, e por proporcionar momentos de risadas, conversas, conselhos, festas e despedidas. Obrigada Raquel, Dani, Hyago e Juliana! Agradeço aos amigos biólogos de Juiz de Fora, pelos bons momentos vividos, pela amizade de sempre e pelos conselhos de vida. Em especial, agradeço à Ana Luiza e Victória, companheiras de pós-graduação, que foram importantes durante o desenvolvimento deste doutorado.

Finalmente, agradeço à minha família, em especial à minha mãe, Fátima, à minha tia, Lurdinha, ao meu pai Ivan e ao meu tio, Ivan pelo apoio, pelo amor dedicado e pelo carinho mesmo estando longe. À minha prima, que considero como irmã, Miriã, e marido, agora primo, Conrado, por acompanhar de perto o meu sonho em Juiz de Fora, e por estar sempre ao meu lado quando preciso.

*“A ciência serve para nos dar uma ideia de quão
extensa é a nossa ignorância.”*

Félicité Robert de Lamennais, 1821

RESUMO

A lignina é uma molécula que confere sustentação e resistência às plantas vasculares. Em plantas forrageiras, como *Urochloa humidicola* e *Cenchrus purpureus*, a molécula interfere negativamente na digestibilidade da fibra vegetal no rúmen de bovinos. Na produção de biocombustíveis, a lignina presente na biomassa dificulta a hidrólise enzimática; na combustão da biomassa, contribui para o aumento do poder calorífico. Dada sua importância, faz-se necessário entender o funcionamento da via metabólica associada a lignina para seleção de genótipos superiores. O objetivo deste trabalho foi identificar genes relacionados à produção de lignina em *U. humidicola* e *C. purpureus*, além do desenvolvimento de marcadores moleculares microsatélites. Neste estudo foram selecionados, para cada espécie, quatro genótipos com maior produção de lignina, e quatro genótipos de menor produção. A análise dos genes diferencialmente expressos em *U. humidicola* revelou 196 genes a nível de significância de $p < 0,05$. Em destaque, genes do citocromo P450, da enzima 4CL e da biossíntese de flavonoides. 53 genes foram diferencialmente expressos em *C. purpureus*, destacando-se os genes glicosiltransferase e os relacionados a biossíntese de terpenos/terpenoides. A partir de um novo agrupamento realizado, foi possível identificar os genes MYB e peroxidase, além de genes de construção da parede celular. Com a construção do *heatmap*, detectou-se maior expressão da família gênica COMT, seguida das famílias CCoAOMT e PTAL. A partir do transcriptoma, foi possível descobrir novos marcadores microsatélites nas duas espécies, que foram testados para amplificação em *U. humidicola* e *C. purpureus*, além de identificação de padrão de bandas únicas em cultivares desta última espécie. Um total de 42 *primers* amplificaram em quatro amostras testadas de *U. humidicola*; 47 amplificaram em quatro amostras testadas de *C. purpureus*, sendo seis informativos na diferenciação de 21 cultivares testadas. Os genes identificados nos transcriptomas serão importantes na identificação dos genes mais promissores para futura manipulação genética. Os marcadores microsatélites desenvolvidos serão utilizados na identificação de cultivares e em estudos de diversidade dentro das populações de melhoramento.

Palavras-chave: RNA-Seq. transcriptoma. genes diferencialmente expressos. lignina. microsatélites.

ABSTRACT

Lignin is a molecule, conferring sustentation and resistance to vascular plants. In forage plants, as *Urochloa humidicola* and *Cenchrus purpureus*, the molecule negatively interferes in vegetal fiber digestibility in bovine rumen. In biofuel production, lignin in biomass is an obstacle to material enzyme hydrolysis; in biomass combustion, lignin helps in calorific values improvement. Given its importance, it is necessary to understand the lignin associated metabolic pathway to superior genotypes selection. This work aims in genes related identification to lignin production in *Urochloa humidicola* and *Cenchrus purpureus*, and specific molecular markers development. We selected, for each species, four genotypes for high lignin production, and four genotypes for low lignin production. Differential expressed genes analysis in *U. humidicola* showed 196 genes to a significant level $p \leq 0.05$. We highlighted cytochrome P450, 4CL and flavonoid biosynthesis genes. 53 genes were differentially expressed in *C. purpureus*, calling attention to glycosyltransferase and terpenes/terpenoids biosynthesis related genes. New analysis using a new grouping showed MYB, peroxidase genes, and genes related to cell wall formation. The heatmap showed COMT gene family as the most expressed gene, followed by CCoAOMT and PTAL families. From transcriptome, it was possible to discover new microsatellite markers in both species, which were tested for amplification in *U. humidicola* and *C. purpureus*, in addition to unique band pattern identification in napiergrass cultivars. 42 primers amplified in 4 *U. humidicola* tested samples; 47 were amplified in 4 *C. purpureus* tested samples, and 6 primers were important in 21 cultivars differentiation. Identified genes from transcriptome will be important in promising genes selection for future genetic manipulation. Developed microsatellite markers will be useful for cultivars identification and breeding population diversity studies.

Key words: RNA-Seq. transcriptome. differentially expressed genes. lignin. microsatellites.

LISTA DE FIGURAS

Figura 1. Estrutura de lignocelulose. Adaptado de Bamdad <i>et al.</i> (2017)	28
Figura 2. Via de biossíntese de lignina. Adaptado de Cesarino <i>et al.</i> (2016)	30
Figura 3. Representação de um polímero de lignina. Adaptado de Stewart <i>et al.</i> (2009)	31
Figura 4: Representação da espécie <i>Urochloa humidicola</i> , cultivar BRS Tupi	37
Figura 5: Representação da espécie <i>Cenchrus purpureus</i>	40
Figura 6. Morfologia de um perfilho em gramíneas. Adaptado de Matthew <i>et al.</i> (2001)	52
Figura 7. Gráfico de Whisker do sequenciamento do transcriptoma dos acessos de <i>Urochloa humidicola</i> BRA000540 (a) e BRA004959 (b). y é o eixo dos valores de qualidades. O fundo do gráfico divide o eixo em três níveis de confiança: muito confiável (verde), qualidade razoável (laranja) e baixa qualidade (vermelho). A linha central azul representa o valor da média da qualidade. A caixa amarela representa a amplitude interquartil (25-75%). Fonte FastQC A Quality Control tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ por S. Andrews	61
Figura 8. Morfologia típica de gramíneas. Adaptado de Ball <i>et al.</i> (2007)	73
Figura 9. Gráfico de Whisker do sequenciamento do transcriptoma dos acessos de <i>Cenchrus purpureus</i> BAGCE 5 (a) e Pioneiro (b). y é o eixo dos valores de qualidades. O fundo do gráfico divide o eixo em três níveis de confiança: muito confiável (verde), qualidade razoável (laranja) e baixa qualidade (vermelho). A linha central azul representa o valor da média da qualidade. A caixa amarela representa a amplitude inter quartil (25-75%). Fonte: FastQC A Quality Control tool for High Throughput Sequence Data http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ por S. Andrews	82
Figura 10. Nível de expressão gênica de genes importantes na via de biossíntese de lignina nos acessos BAGCE 96, BAGCE 98, BAGCE 5 e Pioneiro. No canto superior direito, pode-se observar o indicador de nível de expressão, em TPM (<i>transcripts per</i>	

<i>million</i>). Quanto mais escuro for o azul, maior é a expressão deste gene	89
Figura 11. Nível de expressão gênica importantes na via de biossíntese de lignina, excluindo-se os genes da família COMT, nos acessos BAGCE 96, BAGCE 98, BAGCE 5 e Pioneiro. No canto superior direito, pode-se observar o indicador de nível de expressão, em TPM (<i>transcripts per million</i>). Quanto mais escuro for o azul, maior é a expressão deste gene	90
Figura 12. Géis de agarose dos <i>primers</i> RNA-Bra 01, RNA-Bra 18, RNA-Bra 30 e RNA-Bra 39 nas amostras 22A, 23A, 25A e 28A de <i>U. humidicola</i> inicialmente testadas	101
Figura 13. Géis de agarose dos <i>primers</i> RNA-CE 21, RNA-CE 27, RNA-CE 44 e RNA-CE 12 nas amostras BAGCE 11, BAGCE50, BAGCE 57 e BAGCE 69 de <i>C. purpureus</i> inicialmente testadas	104
Figura 14. Gel de poliacrilamida do <i>primer</i> RNA-CE 05 nas 21 cultivares/acessos de capim-elefante avaliadas. A primeira e últimas colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado.....	106
Figura 15. Gel de poliacrilamida do <i>primer</i> RNA-CE 44 nas 21 cultivares/acessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado	107
Figura 16. Gel de poliacrilamida do <i>primer</i> RNA-CE 46 nas 21 cultivares/acessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado	108
Figura 17. Gel de poliacrilamida do <i>primer</i> RNA-CE 17 nas 21 cultivares/acessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado	109
Figura 18. Gel de poliacrilamida do <i>primer</i> RNA-CE 43 nas 21 cultivares/acessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado	110

LISTA DE TABELAS

Tabela 1. Lista dos oito genótipos de <i>Urochloa humidicola</i> pertencentes ao Programa de Melhoramento de Forrageiras da Embrapa Gado de Corte utilizados no sequenciamento de RNA	50
Tabela 2. Dados da avaliação bromatológica dos oito genótipos de <i>Urochloa humidicola</i> , com as porcentagens de fibra em detergente ácido (FDA) e lignina, em g/kg MS, e classificação dos genótipos com relação ao teor de lignina	58
Tabela 3. Total de <i>reads</i> obtidas em <i>Urochloa humidicola</i> para cada biblioteca gerada, percentual de CG e total de <i>reads</i> geral	60
Tabela 4. Quantidade e porcentagem de <i>reads</i> sequenciadas de <i>U. humidicola</i> mapeadas utilizando como referência <i>U. ruziziensis</i>	63
Tabela 5. Lista dos vinte acessos de capim-elefante (<i>Cenchrus purpureus</i>) pertencentes ao Banco Ativo de Germoplasma de Capim-Elefante da Embrapa Gado de Leite	72
Tabela 6. Análise bromatológica de oito amostras selecionadas de <i>Cenchrus purpureus</i> , com os valores de fibra em detergente ácido (FDA) e lignina, em g/kg MS, e classificação dos acessos com relação ao teor de lignina	79
Tabela 7. Total de <i>reads</i> obtidas em <i>Cenchrus purpureus</i> para cada biblioteca gerada, percentual de CG e total de <i>reads</i> sequenciadas	81
Tabela 8. Alinhamento do sequenciamento ao genoma de capim-elefante, com o número de <i>reads</i> alinhadas e suas porcentagens	83
Tabela 9. Alinhamento das <i>reads</i> sequenciadas ao transcriptoma montado de capim-elefante, com a porcentagem correspondente a cada acesso analisado	84
Tabela 10. Lista dos 42 marcadores microssatélites que amplificaram nas quatro amostras de <i>Urochloa humidicola</i> , com suas sequências <i>forward</i> e <i>reverse</i> , o motivo de repetição e o tamanho de amplificação	98
Tabela 11: Lista dos 47 marcadores microssatélites que amplificaram nas quatro amostras de <i>Cenchrus purpureus</i> , com suas sequências <i>forward</i> e <i>reverse</i> , o motivo de repetição, o tamanho de amplificação, e a identificação dos <i>primers</i> selecionados para o teste em cultivares	102
Tabela 12. Tamanho de bandas amplificadas nas 21 cultivares/acessos de capim-elefante, considerando os <i>primers</i> mais promissores na identificação de um padrão único nas cultivares BRS Capiáçu, BRS Canará, BRS Kurumi e Pioneiro	111

LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS

μL: microlitro

μM: micromolar

°C: graus Celsius

1G: primeira geração

2G: segunda geração

4CL: “4-couamate-CoA ligase” – 4-Curamato-CoA ligase

AC: “activator” – ativador

BAGCE: Banco Ativo de Germoplasma de Capim-Elefante

C3H: “*p*-Couramate 3-hydroxylase” – *p*-curamato 3-hidroxilase

C4H: “cinnamate 4-hydroxylase” – cinamato 4-hidroxilase

CAD: “cinnamyl alcohol dehydrogenase” – cinamoil álcool desidrogenase

CCoAOMT: “caffeoyl Coenzyme A O-Methyltransferase” – cafeoil Coenzima A O-metiltransferase

CCR: “cinnamoyl-CoA reductase” – cinamoil-CoA redutase

cDNA: “complementary DNA” – DNA complementar

CoA: Coenzima A

COMT: “caffeic acid 3-O-methyltransferase” – ácido cafeico O-metiltransferase

DE: diferencialmente expressos

DNA: “deoxyribonucleic acid” – ácido desoxirribonucleico

dNTP: “deoxyribonucleotide triphosphate” – desoxirribonucleotídeo trifosfato

F5H: “ferulate 5-hydroxylase” - ferulato 5-hidroxilase

FAR-1: “Fatty Acil-CoA Reductase 1” – Acil graxo-CoA reductase 1

FDA: fibra em detergente ácido

FDN: fibra em detergente neutro

FDR: “False Discovery Rate” – Taxa de Descoberta Falsa

FIESP/ICONE: Federação das Indústrias do Estado de São Paulo/Instituto de Estudos do Comércio e Negociações Internacionais

G: guaiacil

g/kg: gramas por kilogramas

GC: guanina/citosina

GO/BLAST: “Gene Ontology/Basic Local Alignment Search Tool” – Ontologia de Genes/ Ferramenta de Busca de Alinhamento Básico Local

GT: “glycosyltransferase” – glicosiltransferase

H: *p*-hidroxifenil

HCT: “hydroxyninamoyl-CoA shikimate/quininate” - *p*-Hidroxicinamil-CoA chiquimato transferase

ITS: “Internal Transcribed Space” – Espaçadores Internos Transcritos

KEGG: “Kyoto Encyclopedia of Genes and Genomes” – Enciclopédia de Genes e Genomas de Kyoto

MAPA: Ministério da Agricultura, Pecuária e Abastecimento

min: minutos

mL: mililitro

mM: milimolar

mRNA: “messenger RNA” – RNA mensageiro

MS: matéria seca

MYB: “myeloblastosis” - mieloblastose

MYB4: “myeloblastosis 4” – mieloblastose 4

N: nitrogênio

NAC: acrônimo para NAM (“no apical meristem” - sem meristema apical), ATAF1/2 (“*Arabidopsis thaliana* Activating Factor 1/2” – fator de ativação *Arabidopsis thaliana* 1/2) e CUC2 (“cup-shaped cotyledon” – cotilédone em formato de xícara)

NBCI/BLAST: “National Center for Biotechnology Information/Basic Local Alignment Search Tool” – Centro Nacional para Informação Biotecnológica/Ferramenta de Busca de Alinhamento Básico Local

ng: nanograma

NIRS: “Near-infrared spectroscopy” – espectroscopia no infravermelho próximo

oligo-dT: “oligo-deoxythymine” – oligo-desoxitimidina

nm: nanômetro

PAE: “Pectin acetylerase” – Pectina acetilesterase

PAL: “phenylalanine ammonia-lyase” – fenilalanina amônia-liase

PASA: “Program to Assemble Spliced Alignments” – Programa para montagem de alinhamentos pós-splicing

pb: pares de bases

PB: proteína bruta

PC: poder calorífico

PCR: "Polymerase Chain Reaction" – Reação em Cadeia de Polimerase

PE: "paired-ends"

Pfam: "Protein Families" – Famílias de Proteínas

Proálcool: Programa Nacional do Álcool

PTAL: "phenylalanine/tyrosine ammonia-lyase" - fenilalanina/tirosina amônia-liase

QTL: "Quantitative Trait Loci" – *Loci* de Característica Quantitativa

RAPD: "Random Amplified Polymorphic DNA" – Amplificação Aleatória de DNA Polimórfico

RFLP: "Restriction Fragment Length Polymorphism" – Polimorfismo de Comprimento de Fragmento de Restrição

RIN: "RNA Integrity Number" – Número de Integridade do RNA

RNA: "ribonucleic acid" – ácido ribonucleico

RNA-Seq: sequenciamento de nova geração de RNA

s: segundos

S: siringil

SHMT: "Serine Hydroxymethyltransferase" – serina hidroximetiltransferase SNP: "Single Nucleotide Polymorphism" – polimorfismo de um único nucleotídeo

SSR: "Simple Sequence Repeat" – Repetição de Sequência Simples

STAR: "Spliced Transcripts Aligment to a Reference" – Alinhamento de Transcritos pós-splicing a Referência

TAL: "tyrosine ammonia-lyase" – tirosina amônia-liase

TmHMM: "Transmembrane Helices Hidden Markov Model" – Hélices transmembrana baseado em modelo escondido de Markov

TPM: "transcripts per million" – transcritos por milhão

TrEMBL/SwissProt: "Translated European Molecular Biology Laboratory/SwissProt" – Laboratório Europeu de Biologia Molecular de Tradução/SwissProt

V: volts

XET: "Xyloglucan endotransglucosylase" – Xiloglucano endotransglicosilase

SUMÁRIO

1 INTRODUÇÃO GERAL	18
2 REVISÃO BIBLIOGRÁFICA	21
2.1 PECUÁRIA BRASILEIRA	21
2.2 QUALIDADE DA FORRAGEM	23
2.3 BIOMASSA ENERGÉTICA	24
2.4 LIGNINA E SUA BIOSÍNTESE	27
2.5 ESTUDO DO TRANSCRIPTOMA PELA TÉCNICA DE RNA-Seq	32
2.6 ANÁLISE COMPUTACIONAL PARA AVALIAÇÃO DA EXPRESSÃO DE GENES DIFERENCIALMENTE EXPRESSOS	34
2.7 CARACTERIZAÇÃO DO GÊNERO <i>Urochloa</i> E DA ESPÉCIE <i>Urochloa</i> <i>humidicola</i>	35
2.8 CARACTERIZAÇÃO DO GÊNERO <i>Cenchrus</i> E DA ESPÉCIE <i>Cenchrus</i> <i>purpureus</i>	38
2.9 MARCADORES MOLECULARES	41
2.10 FINGERPRINT DO DNA	43
3 OBJETIVOS	46
3.1 OBJETIVO GERAL	46
3.2. OBJETIVOS ESPECÍFICOS	46

CAPÍTULO 1: ANÁLISE DA EXPRESSÃO GÊNICA EM *Urochloa humidicola*

1.1 INTRODUÇÃO	48
1.2 MATERIAL E MÉTODOS	50
1.2.1 MATERIAL VEGETAL	50
1.2.2 COLETA DE AMOSTRAS E EXTRAÇÃO DE RNA	51
1.2.3. MONTAGEM DE BIBLIOTECA DE cDNA	52
1.2.4 SEQUENCIAMENTO DE mRNA (RNA-Seq)	53
1.2.5 ANÁLISE DE DADOS	54
1.2.5.1 Controle de Qualidade das bibliotecas	54
1.2.5.2 Montagem <i>de novo</i> do transcriptoma	54

1.2.5.3	Indexação e Mapeamento de sequências	55
1.2.5.4	Quantificação de transcritos	56
1.2.5.5	Identificação de genes diferencialmente expressos	56
1.2.5.6	Anotação dos genes	56
1.2.5.7	Identificação dos genes diferencialmente expressos e suas anotações	57
1.3	RESULTADOS E DISCUSSÃO	58
1.3.1	ANÁLISE BROMATOLÓGICA	58
1.3.2	CONTROLE DE QUALIDADE DAS BIBLIOTECAS E MONTAGEM DE NOVO DO TRANSCRIPTOMA	59
1.3.3	INDEXAÇÃO E MAPEAMENTO DE SEQUÊNCIAS	63
1.3.4	IDENTIFICAÇÃO DOS GENES DIFERENCIALMENTE EXPRESSOS E SUAS FUNÇÕES GÊNICAS	64
1.4	CONCLUSÃO	68

CAPÍTULO 2: ANÁLISE DA EXPRESSÃO GÊNICA EM *Cenchrus purpureus*

2.1	INTRODUÇÃO	70
2.2	MATERIAL E MÉTODOS	72
2.2.1	MATERIAL VEGETAL	72
2.2.2	COLETA DE AMOSTRAS E EXTRAÇÃO DE RNA	73
2.2.3	MONTAGEM DE BIBLIOTECA DE cDNA	74
2.2.4	SEQUENCIAMENTO DE mRNA (RNA-Seq)	75
2.2.5	ANÁLISE DE DADOS	76
2.2.5.1	Controle de Qualidade das bibliotecas	76
2.2.5.2	Montagem do transcriptoma	76
2.2.5.3	Indexação e Mapeamento de sequências	77
2.2.5.4	Quantificação de transcritos	77
2.2.5.5	Identificação de genes diferencialmente expressos	77
2.2.5.6	Anotação dos genes	78
2.2.5.7	Identificação dos genes diferencialmente expressos e suas anotações	78
2.3	RESULTADOS E DISCUSSÃO	79

2.3.1 ANÁLISE BROMATOLÓGICA	79
2.3.2 CONTROLE DE QUALIDADE DAS BIBLIOTECAS E MONTAGEM DO TRANSCRIPTOMA	80
2.3.3 INDEXAÇÃO E MAPEAMENTO DE SEQUÊNCIAS	83
2.3.4 IDENTIFICAÇÃO DOS GENES DIFERENCIALMENTE EXPRESSOS E SUAS FUNÇÕES GÊNICAS	84
2.4 CONCLUSÃO	91

**CAPÍTULO 3: IDENTIFICAÇÃO DE MARCADORES MOLECULARES EM *U.
humidicola* E *C. purpureus***

3.1 INTRODUÇÃO	93
3.2 MATERIAL E MÉTODOS	95
3.2.1 IDENTIFICAÇÃO DE MICROSSATÉLITES E DESENHOS DE <i>PRIMERS</i>	95
3.2.2 TESTE DE AMPLIFICAÇÃO DOS <i>PRIMERS</i> DESENHADOS	95
3.2.3 PAINEL ESPECÍFICO PARA IDENTIFICAÇÃO DE CULTIVARES DE <i>Cenchrus purpureus</i>	96
3.3 RESULTADOS E DISCUSSÃO	98
3.3.1 TESTE DE AMPLIFICAÇÃO	98
3.3.2 PAINEL ESPECÍFICO PARA IDENTIFICAÇÃO DE CULTIVARES DE <i>Cenchrus purpureus</i>	105
3.4 CONCLUSÃO	113

REFERÊNCIAS BIBLIOGRÁFICAS	114
-----------------------------------------	------------

1 INTRODUÇÃO GERAL

A lignina é uma molécula presente nas paredes secundárias de plantas vasculares, e compõe, junto com a celulose e a hemicelulose, a lignocelulose (BIDLACK E DASHEK, 2017; SCHUTYSER *et al.*, 2018), matéria prima mais abundante no planeta (HUBER *et al.*, 2006; SCHUTYSER *et al.*, 2018). Pela sua complexidade química, a lignina tem como característica principal garantir resistência a planta (XU *et al.*, 2009). Esta resistência pode ser a fatores bióticos e abióticos, a degradação bacteriana e fúngica em animais herbívoros (RAFFRENATO *et al.*, 2016), e a degradação enzimática em processos de produção energética (WENG *et al.*, 2008; CHEN E DIXON, 2007; VAN ACKER *et al.*, 2013).

As plantas forrageiras, utilizadas na alimentação de bovinos, foram de grande contribuição para o aumento da produção de carne e leite no Brasil. Isto foi causado principalmente pela seleção e pelo melhoramento dos vegetais, que proporcionaram cultivares com maior produção de matéria seca e de melhor qualidade, sem a necessidade de expansão da área de pastagem (GOMES *et al.*, 2017). Dentre as espécies forrageiras, destaca-se as espécies do gênero *Urochloa*, que são responsáveis por 85% das pastagens brasileiras (DO VALLE *et al.*, 2008; VIGNA *et al.*, 2016a). Porém, um problema ainda encontrado está relacionado com a digestibilidade da fibra vegetal, que ainda não é totalmente eficaz em bovinos (BRITO *et al.*, 2003). A principal causa é a presença de lignina na parede celular das plantas, que impede uma melhor digestão do material ingerido (CHERNEY *et al.*, 1991; RAFFRENATO *et al.*, 2016).

Na produção de biocombustíveis, a biomassa vegetal é a matéria-prima utilizada. As plantas C4, como o capim-elefante (*Cenchrus purpureus*), são as principais fontes, pois produzem grande volume de biomassa em um curto período (TAYLOR *et al.*, 2010). A presença de lignina na parede celular dos vegetais se configura como uma barreira física na utilização da biomassa como matéria prima na produção de etanol de segunda geração (2G) (CHEN *et al.*, 2007; WENG *et al.*, 2008; VAN ACKER *et al.*, 2013). Neste processo de produção, a lignina é o principal fator que impede o acesso de enzimas aos polissacarídeos presentes nas paredes

celulares das plantas (WENG *et al.*, 2008; CHEN E DIXON, 2007; VAN ACKER *et al.*, 2013; VICENTINI *et al.*, 2015). Já na produção de energia térmica, a lignina auxilia no aumento do poder calorífico da biomassa (DOS SANTOS *et al.*, 2011).

Dada a importância da lignina e seus diversos papéis biológicos, é fundamental compreender os processos bioquímicos e os genes envolvidos em sua produção, a fim de permitir a manipulação a nível genético do teor de lignina em espécies vegetais. Para espécies com o objetivo de ser utilizada na alimentação animal, como as espécies do gênero *Urochloa* e *Cenchrus*, uma cultivar que apresente um baixo teor de lignina é ideal para melhorar a eficiência da digestão (RAFFRENATO *et al.*, 2016). *Cenchrus purpureus*, além de ser utilizado como forrageira, tem grande potencial para produção de energia térmica e biocombustível. Para melhor eficiência de cada um desses processos, um teor de lignina é desejado: na produção de carvão vegetal, um maior teor de lignina contribui para o aumento do poder calorífico (FERNANDES *et al.*, 2020); na produção de etanol 2G, baixos teores de lignina facilitam o pré-tratamento, o que auxilia na etapa seguinte de hidrólise enzimática (HIMMEL *et al.*, 2007; LI *et al.*, 2016a; ZENG *et al.*, 2020).

O sequenciamento do transcriptoma é uma ferramenta que permite identificar genes expressos e fornece informações sobre as vias metabólicas envolvidas nos processos de interesse. A partir desse conhecimento, é possível identificar quais genes tem atuação predominante dentro da via e fornecerá conhecimento para futuras manipulações genéticas associadas a produção de lignina e busca por marcadores moleculares.

A descoberta de marcadores moleculares, em especial marcadores microsatélites, (TAHERI *et al.*, 2018; LIU *et al.*, 2020a), é uma ferramenta importante em diferentes estudos, como análise de fluxo gênico (LIU *et al.*, 2020a; LU *et al.*, 2021), análise da diversidade gênica (WU *et al.*, 2013; NADEEM *et al.*, 2018) e *fingerprint* de DNA (BANG *et al.*, 2011; HWANG *et al.*, 2020). No mercado de forrageiras, o *fingerprint* de DNA pode ser usado como ferramenta no combate a biopirataria, a partir da investigação da pureza de sementes (JAMIL *et al.*, 2020a; IQBAL *et al.*, 2021). A seleção de marcadores microsatélites em *U. humidicola* e *C. purpureus* permitirá a identificação a nível molecular das cultivares desenvolvidas, confirmando assim a qualidade da cultivar comercializada.

Este trabalho, no geral, visa agregar conhecimento ao programa de melhoramento de ambas as espécies. A partir do RNA-Seq, será possível identificar genes associados a biossíntese de lignina, assim como a sua diferente expressão nas amostras selecionadas a partir do teor de lignina avaliado de *Urochloa humidicola* e *Cenchrus purpureus*. O sequenciamento de RNA também permitirá que novos marcadores moleculares, como microssatélites, sejam desenhados para cada espécie. Em *Cenchrus purpureus*, os microssatélites serão de grande importância na caracterização molecular de quatro cultivares de grande comercialização no mercado de forrageiras, como uma tentativa de garantir a origem de cada cultivar.

2 REVISÃO BIBLIOGRÁFICA

2.1 PECUÁRIA BRASILEIRA

De acordo com o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), em 2021, a projeção do Valor da Produção Pecuária no Brasil será de 330,1 bilhões de reais, um aumento de 5,12% em relação a 2020. Deste valor, a carne bovina representa 45%, superando os valores de carne de frango e produção de leite (<https://www.gov.br/pt-br/noticias/agricultura-e-pecuaria/2021/04/valor-da-producao-agropecuaria-de-2021-deve-ser-o-maior-desde-1989>).

Em 2018, o rebanho bovino brasileiro foi estimado em 214,69 milhões de cabeças, posicionando-o mais uma vez como o maior rebanho bovino do mundo, na frente de países como a Índia (186 milhões), EUA (94,3 milhões) e China (81,5 milhões). Além disso, foi registrado um crescimento de 6,9% no número de abates, chegando a 44,23 milhões de cabeças (ABIEC, 2019).

Na produção de leite bovino, o Brasil se configurou, no ano de 2018, como o quinto maior produtor de leite no mundo, de acordo com dados da Our World in Data (<https://ourworldindata.org>). Neste período, produziu mais de 34 bilhões de litros, ficando atrás de países como Índia, Estados Unidos, Paquistão e China. Na última década, a produção leiteira no Brasil aumentou 32,4% e, ao mesmo tempo, foi observada uma queda no número de vacas ordenhadas, passando de 20.943 milhões de cabeças para 18.606 milhões de cabeças (EMBRAPA, 2018). Estima-se um crescimento bastante significativo na produção de leite nos próximos 10 anos, passando de 36,3 bilhões de litros em 2021 para 43,9 bilhões em 2031 (MAPA, 2021).

Os aumentos expressivos de produção na pecuária podem ser vistos quando se compara a década de 1990 com os últimos anos. De 1990 a 2018, foi observado um aumento de 176% da produtividade. Além disso, a produção de carne também cresceu, aumentando cerca de 140% neste mesmo intervalo de tempo (ABIEC, 2019). A previsão é que o aumento continue, segundo dados do Outlook Brasil 2022, realizado pela FIESP/ICONE. Daqui a 3 anos, espera-se que o rebanho

bovino alcance a marca de 226,9 milhões de cabeças, com produção de carne bovina de 12 milhões de toneladas, e o consumo *per capita* brasileiro de 43,2 kg/habitante/ano.

Os resultados obtidos na cadeia produtiva são graças a sua organização e avanços tecnológicos dos sistemas de produção, dobrando o efetivo do rebanho nas últimas quatro décadas. Alguns elementos importantes, como aumento do ganho de peso dos animais, diminuição da mortalidade, aumento das taxas de natalidade e expressiva diminuição na idade ao abate, foram possíveis devido a adoção de tecnologias pelos produtores rurais nos eixos de alimentação, genética, manejo e saúde animal (GOMES *et al.*, 2017). Outro setor que merece destaque está relacionado com a área de pastagens, que gradualmente sofre uma redução. No ano de 2018, foi estimada em 162,19 milhões de hectares, uma queda de 1,6% comparada ao ano anterior. Nas últimas décadas, esta redução foi de 15% (ABIEC, 2019).

As plantas que constituem as pastagens, ou plantas forrageiras, são aquelas consumidas por animais, em geral ruminantes, e abrangem uma variedade de gêneros e espécies, desde herbáceas até arbustivas (PEREIRA *et al.*, 2001; DO VALLE *et al.*, 2009a). As gramíneas africanas, como as dos gêneros *Panicum*, *Urochloa* e *Cenchrus*, são as mais utilizadas na formação de pastagens no mundo tropical (DO VALLE *et al.*, 2009a).

Além da seleção de espécies com as características desejadas, outro fator que auxiliou no progresso da pecuária no Brasil foi o desenvolvimento de cultivares. Em meados da década de 1980, o melhoramento de forrageiras se baseou na exploração da variabilidade natural de coleções (MILES *et al.*, 1996; HACKER E JANK, 1998; PEREIRA *et al.*, 2001; JANK *et al.*, 2005; MILES, 2007; DO VALLE *et al.*, 2009a). Porém, este processo é finito, pois baseia-se somente na avaliação da capacidade adaptativa de materiais coletados na natureza. O melhoramento de plantas forrageiras por recombinação genética tornou-se uma alternativa para geração de cultivares, visando o aumento da qualidade e produtividade, resistência a pragas e doenças, produção de sementes de boa qualidade, uso suficiente de fertilizantes, adaptação a estresses edáficos e climáticos (DO VALLE *et al.*, 2009a; DO VALLE *et al.*, 2009b) e maior eficiência na sua transformação em produto animal (DO VALLE *et al.*, 2009a).

2.2 QUALIDADE DA FORRAGEM

Uma alimentação animal adequada traz reflexos na qualidade e volume da produção de leite, ganho de peso, eficiência na reprodução (REIS *et al.*, 2016) e na ruminção do gado (HERRERO *et al.*, 2015; LEE *et al.*, 2017). Uma forrageira ideal está relacionada a diversos fatores, como a palatabilidade do material, o tamanho da partícula que é consumida e a qualidade da forragem utilizada (CHERNEY E HALL, 2008).

A qualidade de uma forrageira é a soma dos constituintes das plantas (CHERNEY E HALL, 2008) relacionadas ao valor nutritivo, como a composição química (proteína, fibras, minerais e vitaminas) (ADESOGAN *et al.*, 2017) e a digestibilidade (REIS *et al.*, 2016; WEISS E HALL, 2020). A avaliação do valor nutricional é realizada a partir da estimativa de alguns elementos, como a proteína bruta (PB), a fibra em detergente neutro (FDN) e fibra em detergente ácido (FDA).

A proteína bruta está relacionada principalmente a quantidade de nitrogênio (N) encontrado nas plantas. A avaliação de FDN e FDA é feita a partir do sistema de detergente desenvolvido por Van Soest *et al.* (1991), que divide os alimentos em frações para prever seus valores nutritivos (VAN SOEST, 1994; CHERNEY E HALL, 2008; WEISS E HALL, 2020). A fibra em detergente neutro representa o total de fibras, incluindo celulose, hemicelulose e lignina, que formam as paredes celulares das células vegetais (NEWMAN *et al.*, 2007). Além disso, pode conter concentrações de cinzas e compostos nitrogenados (WEISS E HALL, 2020). Os valores de FDN podem variar de 10% em grãos de milho a 80% em palhas de gramíneas de estações quentes. A fibra em detergente ácido é composta pela porção fibrosa de baixa digestibilidade da FDN (CHERNEY E HALL, 2008), que é a celulose, lignina e sílica (NEWMAN *et al.*, 2007). Estes valores variam de 3% em grãos de milho a 50% em palha de gramíneas de estação quente (NEWMAN *et al.*, 2007).

Uma forrageira considerada de ótima qualidade é aquela que possui uma grande proporção de compostos que são de rápida fermentação, como açúcares, ácidos orgânicos e proteínas, e baixa proporção de fibras (WAGHORN E CLARK, 2004; LEE *et al.*, 2017), que são de difícil digestão. Por isso, quanto maior o valor de

PB e menores os valores de FDN e FDA, melhor é a forrageira nos quesitos de qualidades de ingestão e digestão (NEWMAN *et al.*, 2007).

As gramíneas encontradas no Brasil, as de origem tropical, apresentam, em comparação às gramíneas de clima temperado, maior porcentagem de FDN e FDA (média de 21% a mais) e menor quantidade de PB (média de 8% a menos) (LEE *et al.*, 2017). A redução do valor nutritivo está relacionada a mecanismos de prevenção aos estresses térmico e hídrico, como maior proporção de caule:folha, alta proporção de tecidos vasculares, e alta concentração de lignina e sílica (KERING *et al.*, 2011; LEE *et al.*, 2017).

Além do fator climático, a maturidade das plantas é um fator que contribui para uma maior concentração de fibras. Em plantas maduras, a proporção de caule em relação a folha aumenta (REIS *et al.*, 2016), o que acarreta maior fração de fibras, menor proporção de PB e menor digestibilidade (NEWMAN *et al.*, 2007; REIS *et al.*, 2016).

O processo de digestão no rúmen é dependente da população microbiana que ali se encontra, pois as bactérias são responsáveis pela degradação da fibra da forrageira (NEWMAN *et al.*, 2007). Porém, a lignina encontrada nas fibras é resistente a degradação bacteriana e fúngica (RAFFRENATO *et al.*, 2016), e as interações químicas e físicas da lignina com as porções de carboidratos na parede celular prejudicam a digestão dos mesmos (CHERNEY *et al.*, 1991; RAFFRENATO *et al.*, 2016).

2.3 BIOMASSA ENERGÉTICA

Com a crescente demanda de energia, aliada ao desenvolvimento da economia mundial (SU *et al.*, 2015), a busca por fontes alternativas e renováveis se torna cada vez mais necessária. Uma dessas fontes é o biocombustível, energia produzida a partir de plantas ou de resíduos da agricultura, e que pode contribuir para a redução da dependência mundial em diesel ou gasolina. Em comparação aos combustíveis fósseis, essa fonte é renovável e mais barata, emitindo até 4 vezes menos gases do efeito estufa (BORIN *et al.*, 2017).

Os tipos principais de biocombustíveis são aqueles chamados de primeira geração (1G), como o biodiesel, biogás e bioetanol, sendo o último o mais usado mundialmente (POPP *et al.*, 2016; BORIN *et al.*, 2017). No ano de 2014, a produção de bioetanol alcançou a marca de 24 milhões de galões, sendo os maiores produtores os Estados Unidos e o Brasil. Juntos, produziram mais de 80% da produção global (BORIN *et al.*, 2017).

O Brasil é considerado atualmente o segundo maior produtor e consumidor do mundo, ficando atrás somente dos Estados Unidos (CARNEIRO *et al.*, 2017; CASTILLO *et al.*, 2017). A produção de etanol no país foi impulsionada pela crise do óleo na década de 1970, por meio de uma série de medidas para melhorar a produção de etanol (SU *et al.*, 2015). O Programa Nacional do Álcool, o Próalcohol, foi iniciado focando na produção de etanol, a partir da fermentação e cristalização do açúcar encontrado no caule de cana-de-açúcar (ARRUDA, 2012; VICENTINI *et al.*, 2015). O sucesso do programa destacou o Brasil como o melhor modelo econômico na indústria de etanol, graças a sua tecnologia e sua vasta quantidade de terra disponível para cultura da cana-de-açúcar (CANILHA *et al.*, 2012; DELLA-BIANCA *et al.*, 2013; BORIN *et al.*, 2017). No ano de 2008, o etanol de cana-de-açúcar produzido no país representou cerca de 37,3% da produção mundial, de 65,6 bilhões de litros (PESTANA-CALSA *et al.*, 2012). Entretanto, a produção de bioetanol de primeira geração compete com o mercado alimentício, inflando o preço dos alimentos (WAI *et al.*, 2017). O etanol de segunda geração (2G), que utiliza como matéria-prima a biomassa vegetal (VICENTINI *et al.*, 2015), torna-se uma alternativa para diminuir a competição e aumentar a segurança alimentar (LOVETT *et al.*, 2009; VALENTINE *et al.*, 2012; LI E CHEN, 2018).

A biomassa vegetal utilizada pode ser caracterizada como toda matéria orgânica de origem vegetal, que pode ser convertida em energia. É considerada uma fonte indireta da energia solar, pois ela é convertida em energia química, através da fotossíntese. Esta matéria pode ser utilizada na produção de energia elétrica, pela combustão direta, energia térmica, como carvão vegetal e lenha, e energia mecânica, como os biocombustíveis de segunda geração (ANEEL, 2005). As espécies ideais para a produção dessas energias são aquelas que produzem um grande volume de biomassa em um curto período de tempo, como as plantas C4, representadas por

espécies de *Miscanthus*, *Panicum virgatum* (LEWANDOWSKI *et al.*, 2003; VAN DER WEIJDE *et al.*, 2013) e *Cenchrus purpureus* (RUEDA *et al.*, 2016).

Na produção de etanol 2G, o processamento é feito a partir da hidrólise enzimática da celulose, que é convertida em açúcares que serão posteriormente fermentados para produção do álcool (YUAN *et al.*, 2014; VICENTINI *et al.*, 2015). Porém, polímeros de lignina e hemicelulose interferem na eficácia das enzimas envolvidas no processo de produção de etanol, um processo conhecido como recalitrância da biomassa (WENG *et al.*, 2008; CHEN E DIXON, 2007; VAN ACKER *et al.*, 2013; VICENTINI *et al.*, 2015; HIMMEL *et al.*, 2007; LI *et al.*, 2016, ZENG *et al.*, 2020). Sendo assim, é necessário que a biomassa passe pelo processo de sacarificação, um pré-tratamento físico e/ou químico que facilita o acesso da enzima que quebra os complexos polissacarídeos (BORIN *et al.*, 2017).

Para a produção de energia térmica, alguns critérios são utilizados para classificar uma biomassa mais eficiente. No processo de combustão, é desejado que o teor de umidade esteja em torno de 50% (MCKENDRY, 2002; VIEIRA, 2012; FERNANDES *et al.*, 2020) e que as concentrações de cinzas e de nitrogênio sejam baixas (cerca de 5% e 1%, respectivamente) (MCKENDRY, 2002; LALAK-KANCZUGOWSKA *et al.*, 2016; FERNANDES *et al.*, 2020), pois reduzem o valor calorífico da biomassa (PROTÁSIO *et al.*, 2011; FERNANDES *et al.*, 2020). Em contrapartida, altos valores de poder calorífico (PC), carbono e hidrogênio são desejados (MCKENDRY, 2002; JARADAT, 2010; FERNANDES *et al.*, 2020). O PC está relacionado a energia liberada quando a biomassa está sendo queimada (FERNANDES *et al.*, 2020), e um maior PC torna uma matéria-prima de eficiente combustão (JARADAT, 2010; FERNANDES *et al.*, 2020). O valor energético de carbono e hidrogênio tem uma correlação positiva com a liberação de energia na combustão, e as principais fontes destas moléculas são a celulose e a lignina (FERNANDES *et al.*, 2020). Por isso, são realizadas análises químicas destes componentes da parede celular, a fim de se caracterizar uma biomassa. Na biomassa de coníferas, importantes matérias primas na combustão, pode-se encontrar valores, em média, de 35-40% de celulose e 27-30% de lignina (MCKENDRY, 2002).

A seleção e o melhoramento de espécies vegetais que atendam às necessidades da produção de energia são etapas importantes para uma melhor eficiência neste setor. Plantas com teor de lignina mínimo para a sustentação das

plantas, e que permitam o acesso de enzimas no processo de sacarificação, são as ideais na produção de etanol 2G. Por outro lado, espécies que apresentam um alto teor de fibras, lignina e celulose são escolhas certas na produção de energia térmica, para uma eficiente queima da biomassa vegetal.

2.4 LIGNINA E SUA BIOSÍNTESE

A lignina é o segundo polímero mais abundante na terra (BAJWA *et al.*, 2019), representando cerca de 30% da biomassa total produzida na biosfera (BOERJAN *et al.*, 2003). Ela está envolvida em funções como resistência mecânica, dando força para a estatura ereta das plantas (HIGUCHI, 2006; YU *et al.*, 2020); transporte de nutrientes, água e metabólitos a longa distância (BAUCHER *et al.*, 2003; BONAWITZ E CHAPPLE, 2010); proteção contra ação de patógenos e resistência a estresses bióticos e abióticos (QUENTIN *et al.*, 2009), como a resistência a seca (LIU *et al.*, 2018; XU *et al.*, 2020).

Nas plantas, a lignina está presente na parede celular secundária, uma camada encontrada internamente à parede celular primária de tipos celulares especializados, como vasos de xilema e fibras de esclerênquima (COSGROVE E JARVIS, 2012; CESARINO *et al.*, 2016). O polímero, junto à celulose e hemicelulose, forma a lignocelulose (Figura 1), o principal componente estrutural em plantas (BIDLACK E DASHEK, 2017) e o tipo mais abundante de biomassa terrestre (HUBER *et al.*, 2006).

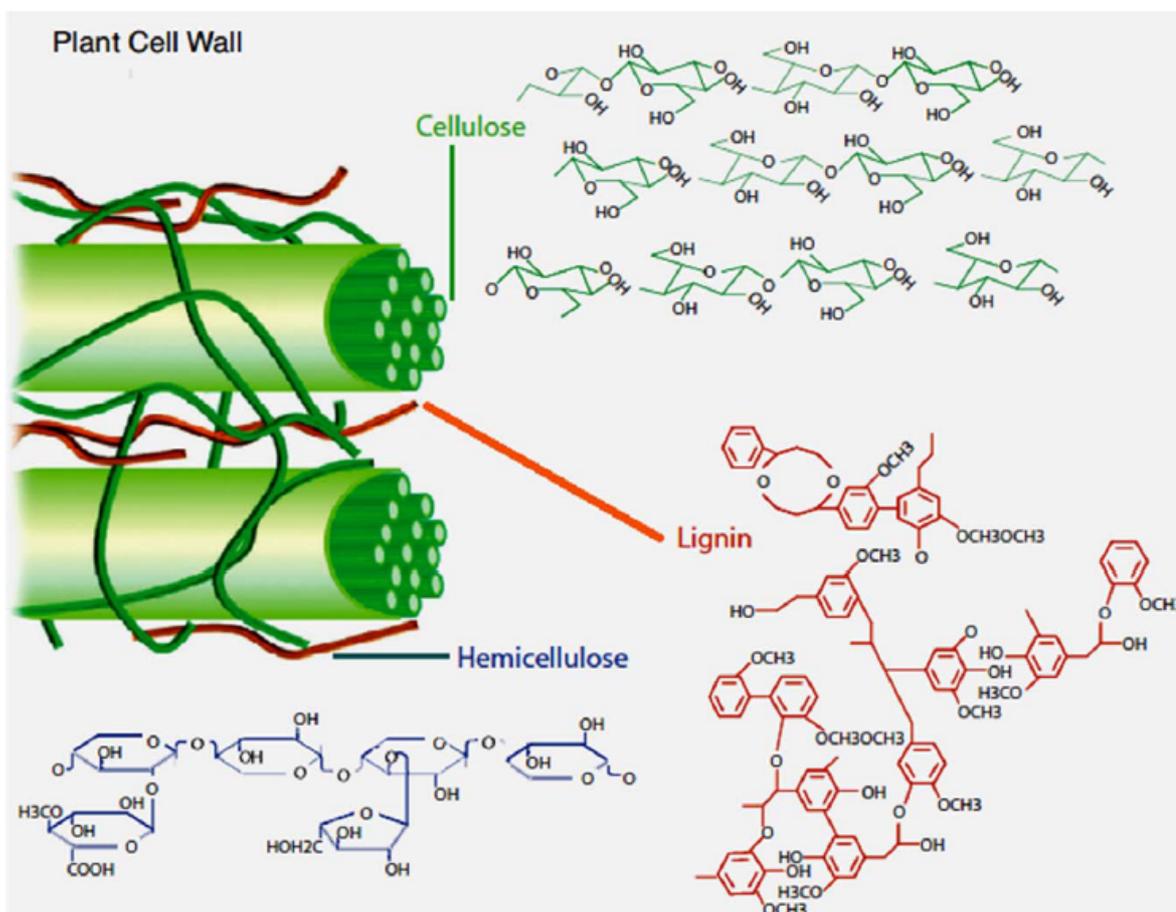


Figura 1: Estrutura de lignocelulose. Adaptado de Bamdad *et al.* (2017).

A lignina é um polímero aromático heterogêneo (BONAWITZ E CHAPPLE, 2010), formado por monômeros derivados majoritariamente da via do fenilpropanoide, cujo precursor primário é o aminoácido aromático fenilalanina (CESARINO *et al.*, 2016). A via é altamente conservada em plantas vasculares, e a maioria das enzimas envolvidas na biossíntese já foram identificadas e caracterizadas (BOERJAN *et al.*, 2003; XU *et al.*, 2009). A via é iniciada com a desaminação da fenilalanina, catalisada pela enzima PAL (fenilalanina amônia-liase) (Figura 2), produzindo ácido cinâmico, a partir do qual compostos fenólicos são produzidos em vegetais superiores (CAROCHA *et al.*, 2015). Em seguida, o ácido cinâmico é convertido em ácido *p*-cumárico, catalisado pela enzima C4H (cinamato 4-hidroxilase) (Figura 2), uma monoxidase dependente de citocromo P450 (SEWALT *et al.*, 1997; BONAWITZ E CHAPPLE, 2010).

Estudos da via em monocotiledôneas relatam que a biossíntese de lignina pode ser iniciada alternativamente a partir do aminoácido tirosina (DEL RÍO *et al.*, 2012 a,b). Nesta via, a tirosina é desaminada pela ação da enzima TAL (tirosina

amônia-liase) (Figura 2) ou PTAL (fenilalanina/tirosina amônia-liase) (BARROS *et al.*, 2016), convertendo o aminoácido em *p*-cumárico. Esta via ignora a reação catalisada pela enzima C4H, presente na via do fenilpropanoide (ROSLER *et al.*, 1997; WATTS *et al.*, 2006).

Em ambas as vias, a biossíntese prossegue pela ação da enzima 4CL (4-Curamato CoA ligase) (Figura 2), envolvida na formação de *p*-cumaroil-coenzima A (*p*-cumaroil-CoA) a partir do ácido *p*-cumárico (EHLTING *et al.*, 2001; BONAWITZ E CHAPPLE, 2010). A molécula produzida está envolvida em diferentes rotas metabólicas, como na produção de flavonoides, *p*-curamatos, ésteres, taninos e estilbenos (BONAWITZ E CHAPPLE, 2010).

Na via de biossíntese dos monolignóis de lignina, o *p*-cumaroil-CoA é utilizado como substrato por duas enzimas diferentes: CCR (cinamoil-CoA redutase) e HCT (*p*-Hidroxicinamil-CoA chiquimato transferase) (BONAWITZ E CHAPPLE, 2010).

A enzima CCR catalisa a produção de *p*-cumaraldeído (Figura 2), que é substrato da reação catalisada pela enzima CAD (cinamil álcool desidrogenase) (BONAWITZ E CHAPPLE, 2010). O produto gerado, o álcool *p*-cumaril, é o monolignol precursor da unidade *p*-hidroxifenil (H) (Figura 2) (BOERJAN *et al.*, 2003).

A enzima HCT catalisa duas reações na via de biossíntese: inicialmente, substitui em *p*-cumaroil-CoA a ligação CoA para ligação chiquimato, formando *p*-cumaroil chiquimato (Figura 2). A enzima C3H (*p*-curamato 3-Hidroxilase), uma monoxidase dependente de citocromo P450, adiciona um grupamento hidroxila ao benzeno de *p*-cumariol chiquimato, formando o cafeoil-chiquimato (Figura 2). Esta molécula é substrato para a segunda reação que envolve HCT, que substitui a ligação chiquimato para ligação CoA, produzindo cafeoil-CoA (BONAWITZ E CHAPPLE, 2010; CESARINO *et al.*, 2016).

Seguindo a via, o cafeoil-CoA sofre metilação catalisada pela enzima CCoAOMT (cafeoil-CoA O-metiltransferase), formando a molécula feruloil-CoA (Figura 2). O feruloil-CoA sofre reação catalisada por CCR (cinamoil-CoA redutase), perdendo a ligação CoA, e produzindo a molécula coniferaldeído (Figura 2) (BONAWITZ E CHAPPLE, 2010; CAROCHA *et al.*, 2015; CESARINO *et al.*, 2016).

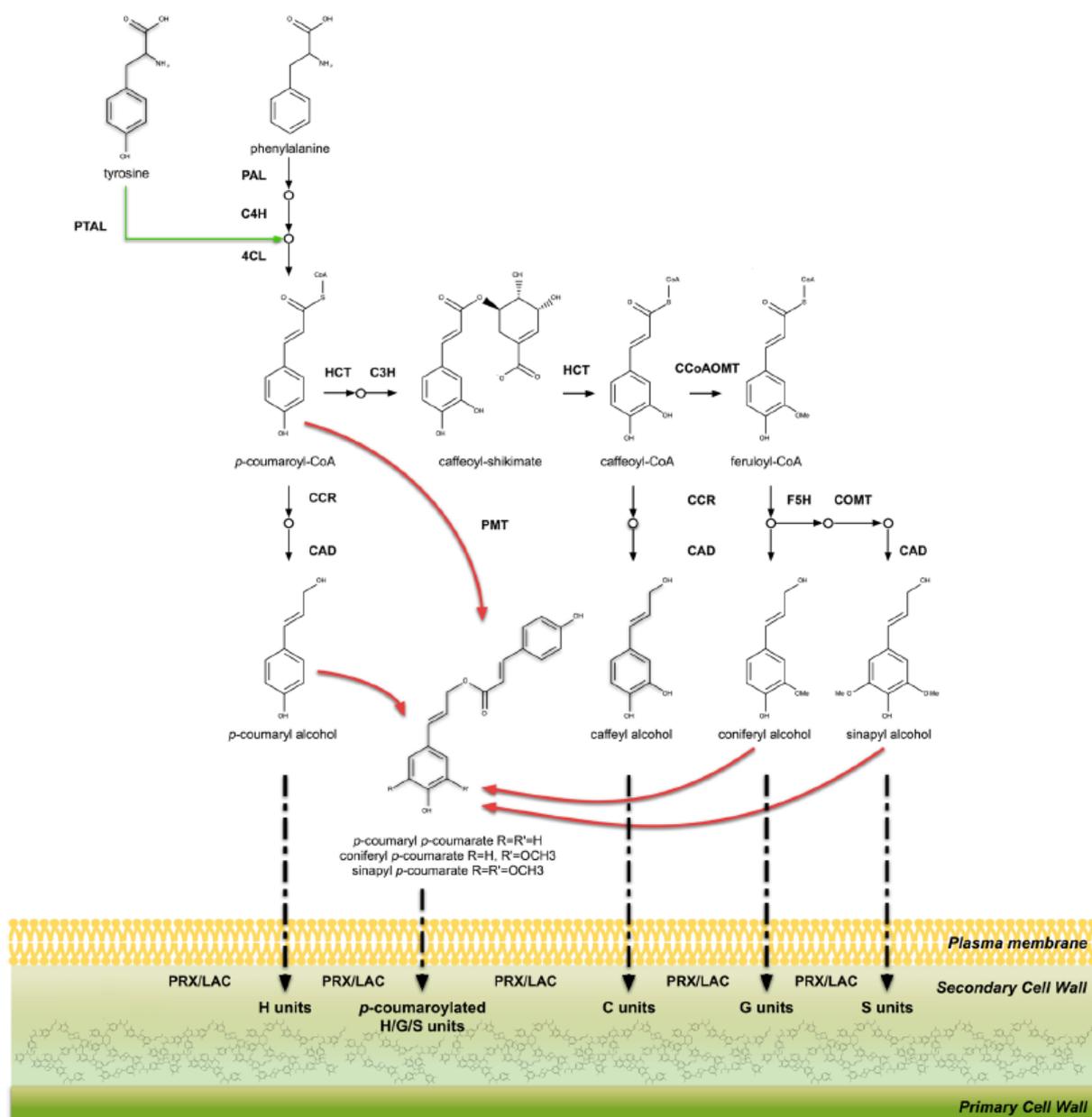


Figura 2: Via da biossíntese de lignina. Adaptado de Cesarino *et al.* (2016).

A molécula coniferaldeído é precursor dos dois tipos de monolignóis que dão origem as unidades de maior presença nas paredes secundárias celulares: unidades guaiacil (G) e siringil (S) (BONAWITZ E CHAPPLE, 2010). Para a produção de álcool coniferil, álcool hidroxicinamol que dá origem a unidade G, o coniferaldeído é reduzido a álcool, utilizando como agente catalisador a enzima CAD (cinamoil álcool desidrogenase) (Figura 2) (SIBOUT *et al.*, 2003; BONAWITZ E CHAPPLE, 2010). Para a produção de álcool sinapil, que dá origem a unidade S, o coniferaldeído ganha uma hidroxila na posição 5' do benzeno, catalisado pela enzima F5H (ferulato 5-

hidroxilase). Em angiospermas, essa enzima é o principal fator responsável pela alteração da razão das unidades G:S (RALPH *et al.*, 2019). Ao final, é produzido 5-hidroxiciniferaldeído, o qual é metilado pela enzima COMT (ácido cafeico O-metiltransferase), formando sinapaldeído. Esta molécula é reduzida a álcool, catalisado por CAD, produzindo o monolignol álcool sinapil (Figura 2) (BONAWITZ E CHAPPLE, 2010).

Após a sua síntese, os monolignóis são transportados para a parede celular secundária, onde são oxidados e polimerizados (Figura 2) (BOERJAN *et al.*, 2003). A formação da lignina se dá pela polimerização hidrogenativa dos monolignóis transportados (CHRISTENSEN *et al.*, 2000), catalisada por diferentes classes de proteínas, como peroxidases, lacases, polifenol oxidases e álcool coniferil oxidases (BOERJAN *et al.*, 2003), que agem de forma individual ou em conjunto (ÖNNERUD *et al.*, 2002), e que ao final gera radicais de monolignóis (BOERJAN *et al.*, 2003). A origem dos radicais é conhecida por meio de transferência de radicais, tanto de monolignóis quanto de outros intermediários, como o *p*-curamato. Esta molécula está relacionada principalmente com a produção de radicais do álcool sinapil (TAKAHAMA E ONIKI, 1994; HATFIELD *et al.*, 1997). Em seguida, os radicais são acoplados ao polímero de lignina em crescimento, processo conhecido como lignificação (Figura 3) (BOERJAN *et al.*, 2003).

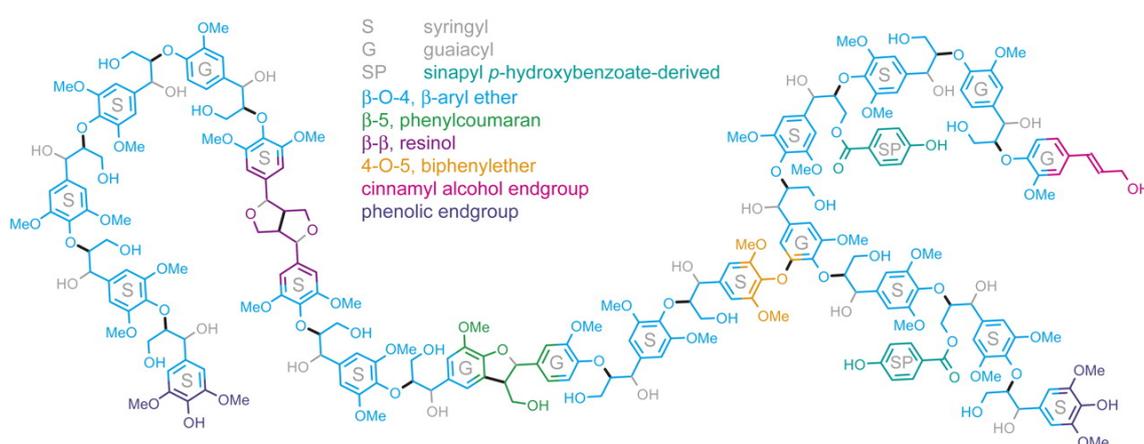


Figura 3: Representação de um polímero de lignina. Adaptado de Stewart *et al.* (2009).

A lignificação em plantas demanda um alto consumo de energia (GENG *et al.*, 2020), e é considerado um processo irreversível (OHTANI E DEMURA, 2019;

GENG *et al.*, 2020). Por isso, torna-se necessário que o tempo e o lugar de depósito de lignina sejam controlados (OHTANI E DEMURA, 2019).

Alguns mecanismos podem ser vistos para um efetivo controle da produção da molécula, como a variação da quantidade de lignina entre tecidos e tipos celulares (VANHOLME *et al.*, 2019; OLIVEIRA *et al.*, 2020), entre diferentes estágios de desenvolvimento (BARROS *et al.*, 2015; OLIVEIRA *et al.*, 2020), e fatores de transcrição que auxiliam na regulação da biossíntese de lignina (VANHOLME *et al.*, 2010; OHTANI E DEMURA, 2019).

Os principais fatores de transcrição descritos na regulação da biossíntese de lignina são os da família do NAC e do MYB (ZHONG E YE, 2009; WANG E DIXON, 2012; NAKAMO *et al.*, 2015; YU *et al.*, 2020), que regulam por meio de uma rede de diversos níveis (ZHAO E DIXON, 2011; OHTANI E DEMURA, 2019; XU *et al.*, 2020) os promotores dos genes da biossíntese de lignina a partir dos elementos AC contidos na sequência (RAES *et al.*, 2003; GENG *et al.*, 2020).

2.5 ESTUDO DO TRANSCRIPTOMA PELA TÉCNICA DE RNA-Seq

O Sequenciamento do Transcriptoma, conhecido como RNA-Seq, é uma tecnologia aplicada em diversos estudos como identificação de variações na sequência como SNPs (*Single Nucleotide Polymorphism*) (CLOONAN *et al.*, 2008; MORIN *et al.*, 2008; WANG *et al.*, 2009), que podem ser usados para detecção de sequências curtas com modificações pós-transcricionais (MARGUERAT E BÄHLER, 2010; CARDOSO-SILVA *et al.*, 2014). Também permite revelar, a regulação da expressão de genes por RNAs não codificantes (DJEBALI *et al.*, 2012; MORRIS E MATTICK, 2014), RNAs potenciadores (LI *et al.*, 2016b) e, em destaque, simultaneamente novos transcritos e suas isoformas (CARDOSO-SILVA *et al.*, 2014) e identificar genes diferencialmente expressos entre espécies (MOROZOVA E MARRA, 2008; CARDOSO-SILVA *et al.*, 2014; STARK *et al.*, 2019), tecidos distintos, genótipos, condições abióticas e estágios de desenvolvimento (HUANG *et al.*, 2017; OLIVEIRA *et al.*, 2018). A importância desta técnica em estudos de expressão gênica se deve ao fato de ser mais barata e fácil de implementar comparado a técnica de

microarray, somada a geração de dados mais compreensíveis e de alta qualidade, capturando níveis de expressão quantitativas em todo o transcriptoma (STARK *et al.*, 2019).

A técnica se inicia pela conversão de uma população de RNA, total ou fracionada, para uma biblioteca de fragmentos de cDNA, e adaptadores ligados em uma ou em ambas as extremidades. Cada molécula é então sequenciada em modo de alto rendimento, obtendo sequências curtas de uma extremidade (sequenciamento *single-end*) ou de ambas as extremidades (sequenciamento *paired-end*). As sequências curtas, chamadas de *reads*, variam de 30 a 400 pares de bases, dependendo da tecnologia de sequenciamento de RNA empregada. Após o sequenciamento, as *reads* geradas podem ser alinhadas a um genoma de referência ou transcritos de referência, ou montadas *de novo*, sem o uso de um genoma de referência, e produzem um mapa de transcritos de escala genômica, que consiste na estrutura transcricional e/ou nível de expressão de cada gene (WANG *et al.*, 2009).

Em plantas, a primeira espécie vegetal que obteve o sequenciamento do transcriptoma por RNA-Seq foi *Arabidopsis thaliana*, utilizando a plataforma 454 da Roche (WEBER *et al.*, 2007; SCHLIESKY *et al.*, 2012). Após este trabalho, diversas espécies vegetais como eucalipto, melão e diferentes leguminosas também tiveram seu transcriptoma sequenciado, e os resultados encontrados foram importantes em identificar marcadores para o melhoramento molecular (NOVAES *et al.*, 2008; GUO *et al.*, 2010; BLAVET *et al.*, 2011; HIREMATH *et al.*, 2011; KAUR *et al.*, 2011; SCHLIESKY *et al.*, 2012). Anos após, o sequenciamento de *A. thaliana* pela Roche, a empresa Illumina também iniciou o sequenciamento de transcriptoma, com espécies de genoma já sequenciado (DENYER *et al.*, 2019; FAN *et al.*, 2020; ZAINAL-ABIDIN *et al.*, 2020). A plataforma então foi utilizada para análise de transcriptomas de espécies modelos (FARMER *et al.*, 2021; PARK *et al.*, 2021), e de interesse comercial (KONG *et al.*, 2020; LIU *et al.*, 2020b; CHEN *et al.*, 2021; XU *et al.*, 2021), incluindo espécies da família Poaceae.

Na família Poaceae, já estão depositadas em banco de dados públicos 67 genomas sequenciados, e cerca de 25 milhões de sequências de DNA e RNA (NCBI, 2021). Apesar de apresentar um significativo volume de dados, a maioria das sequências são derivadas de poucas espécies de importância agronômica, como arroz, milho, trigo e sorgo (SALGADO *et al.*, 2017).

2.6 ANÁLISE COMPUTACIONAL PARA AVALIAÇÃO DA EXPRESSÃO DE GENES DIFERENCIALMENTE EXPRESSOS

Em estudos de análise de genes diferencialmente expressos de organismos, são utilizadas ferramentas computacionais que permitem a identificação dos genes envolvidos nas vias biológicas de interesse, além de seus níveis de expressão. A seleção das ferramentas em diferentes etapas da análise leva em consideração a espécie estudada, além de recursos computacionais disponíveis para análise (SAHAEIAN *et al.*, 2017).

De acordo com Stark *et al.* (2019), as análises computacionais realizadas podem ser divididas em quatro etapas: (i) mapeamento das *reads* em um transcriptoma ou genoma de referência; (ii) quantificação das *reads* associadas a cada gene ou transcrito; (iii) normalização das contagens, para filtragem de genes de baixa expressão; (iv) análise estatística confiáveis relacionadas à expressão diferencial.

A primeira fase é o mapeamento das *reads* do sequenciamento a um transcriptoma ou genoma conhecido, convertendo cada *read* em uma ou mais coordenadas genômicas. Diversas ferramentas estão disponíveis, a exemplo ferramenta STAR, que inicialmente gera uma indexação do genoma de referência, e em seguida, é feito o mapeamento das *reads* propriamente dito (DOBIN *et al.*, 2013). Em algumas espécies, não é possível encontrar um transcriptoma ou um genoma sequenciado e montado disponível. Neste caso, as *reads* sequenciadas podem ser usadas em uma etapa de montagem do transcriptoma (TRAPNELL *et al.*, 2010). Ferramentas de montagem como StringTie (PERTEA *et al.*, 2015) utilizam *reads* alinhadas a um genoma montado de uma espécie próxima àquela em estudo para preencher lacunas identificadas nos alinhamentos para derivar os limites dos éxons e possíveis de regiões de *splice*. Outras ferramentas se baseiam na montagem *de novo* do transcriptoma somente nos dados gerados pelo RNA-seq. O *software* Trinity (GRABHERR *et al.*, 2011) gera transcritos que são organizados em grafos de Bruijn, os quais são utilizados para a montagem do transcriptoma final.

A quantificação das *reads* sequenciadas faz parte da segunda etapa das análises computacionais. A quantificação é feita a partir da contagem das *reads*

sequenciadas que sobrepõem genes conhecidos do transcriptoma de referência. Entretanto, alocar *reads* curtas requer uma etapa de estimativa, para evitar que sequências sejam atribuídas de forma ambígua a uma isoforma específica (VIJAY *et al.*, 2013). Também estão disponíveis ferramentas de quantificação que descartam a etapa de alinhamento, como o Salmon (PATRO *et al.*, 2017). As ferramentas têm demonstrado bom desempenho em caracterizar transcritos mais abundantes, porém são menos precisos em quantificar transcritos menos abundantes ou curtos (WU *et al.*, 2018).

Na terceira fase, é possível filtrar e normalizar transcritos de acordo com o padrão de expressão e viés técnico (RISSO *et al.*, 2011; WAGNER *et al.*, 2012; RISSO *et al.*, 2014). Transcritos que apresentem baixa abundância são detectados e filtrados, permitindo a detecção de maior acurácia de verdadeiras expressões diferenciais (BOURGON *et al.*, 2010). Essa etapa geralmente está presente em pacotes de análises estatísticas de diferenciação de expressão, que é a última fase de análise.

A partir da contagem de *reads* obtidas na segunda fase, é feita uma modelagem para determinar quais transcritos presentes têm mais chance de terem mudado seu nível de expressão. Diversas ferramentas podem ser usadas, a exemplo do pacote “DESeq2” (LOVE *et al.*, 2014) encontrado no programa R (<http://www.R-project.org>), que avalia a expressão a nível gênico.

2.7 CARACTERIZAÇÃO DO GÊNERO *Urochloa* E DA ESPÉCIE *Urochloa humidicola*

O gênero *Urochloa* pertence à família Poaceae, tribo Paniceae, e possui cerca de 120 espécies (WATSON E DALLWITZ, 1992; VIGNA *et al.*, 2016a), distribuídas em regiões tropicais e subtropicais dos continentes americanos, asiático, na Oceania e, especialmente, no continente africano (KELLER-GREIN *et al.*, 1996; KARIA *et al.*, 2006).

O gênero sofreu algumas mudanças nas últimas décadas, principalmente pela incorporação de algumas espécies pertencentes ao gênero *Brachiaria*. TORRES GONZÁLEZ E MORTON (2005) estudaram as relações filogenéticas de espécies

identificadas nos gêneros *Urochloa* e *Brachiaria* por meio de análises moleculares da região ITS (*Internal Transcribed Space*) de DNA ribossomal e de características morfológicas. Neste trabalho, o gênero *Brachiaria* foi classificado como um grupo monofilético dentro de *Urochloa*, formando um complexo *Urochloa-Brachiaria*.

Algumas espécies do gênero são conhecidas como gramíneas forrageiras tropicais amplamente cultivadas na América do Sul, Austrália e leste da Ásia (MILES *et al.*, 1996; KUWI *et al.*, 2018). Reúnem características importantes como alta produção de forragem (BOGDAN *et al.*, 1977; KUWI *et al.*, 2018), excelente qualidade nutricional e ampla adaptabilidade (MILES *et al.*, 1996; KUWI *et al.*, 2018). No Brasil, o gênero tornou-se economicamente importante nos últimos 30 anos, viabilizando a atividade pecuária nos solos ácidos e de baixa fertilidade natural do cerrado (DO VALLE *et al.*, 2009a). As espécies de maior destaque no país foram introduzidas da África, e são responsáveis por 85% das pastagens cultivadas no país (DO VALLE *et al.*, 2008; VIGNA *et al.*, 2016a).

Os avanços de produtividade da bovinocultura no Brasil se deram em parte pela contribuição dada pelos programas de melhoramento através da disponibilização de cultivares adaptadas a diferentes regiões (DO VALLE *et al.*, 2008; VIGNA *et al.*, 2016a). Os programas visam a obtenção de cultivares produtivas e persistentes, com boa qualidade nutricional, adaptação a estresses bióticos e abióticos, produção de sementes de boa qualidade, adaptação ao pastejo, assegurando a transformação da pastagem em proteína animal de alto valor (PEREIRA *et al.*, 2001; KARIA *et al.*, 2006; DO VALLE *et al.*, 2009a).

A espécie *Urochloa humidicola* (Rendle) Morrone & Zuloaga (sinonímia *Brachiaria humidicola*) (Figura 4) é muito utilizada em pastagens no Brasil e em outros ambientes tropicais, particularmente em solos ácidos e áreas mal drenadas (KELLER-GREIN *et al.*, 1996; VIGNA *et al.*, 2016a), ou áreas com inundação sazonal (JUNGMANN *et al.*, 2010). *U. humidicola* é adaptada a sistemas de baixo nitrogênio e cresce bem na presença de amônia ou nitrato (CASTILLA E JACKSON, 1991; NUÑEZ *et al.*, 2018), diferente de algumas culturas vegetais que crescem melhor em lugares suplementados com nitrato (MARSCHNER, 1995; RAO *et al.*, 1996; NUÑEZ *et al.*, 2018). Esta espécie tem o centro de origem e dispersão na África oriental (JUNGMANN *et al.*, 2010), e é caracterizada como uma planta perene estolonífera, com caules estreitos e finos, e folhas estreitas e glabrasas (MILES *et al.*, 2004;

JUNGMANN *et al.*, 2010). O modo reprodutivo dominante da espécie é por apomixia apospórica pseudogâmica que, embora facultativa, prevalece sobre o modo sexual (JUNGMANN *et al.*, 2009).



Figura 4: Representação da espécie *Urochloa humidicola*, cultivar BRS Tupi.

U. humidicola é uma espécie poliploide, com número básico cromossômico de $x=6$ (ADAMOWSKI *et al.*, 2007; BOLDRINI *et al.*, 2009a; BOLDRINI *et al.*, 2009b; BOLDRINI *et al.*, 2010; VIGNA *et al.*, 2016b) sendo encontrados em biotipos sexuais indivíduos hexaploides ($2n=6x=36$). Em indivíduos apomíticos, as ploidias encontradas variam de hexa a nonaploides ($2n=9x=54$) (JUNGMANN *et al.*, 2009).

A identificação de *U. humidicola* se confunde a identificação da espécie *Urochloa dictyoneura*. De acordo com Tropicos®, o banco de dados de nomenclatura do Missouri Botanical Garden, *U. humidicola* já foi classificada como *Brachiaria dictyoneura* (Fig. e De Not.) Veldkamp, por Veldkamp (1996) e *Brachiaria dictyoneura* subespécie *humidicola* (Rendle) Catasús, por Catasús Guerra (2001). Atualmente, a espécie recebeu a nomenclatura de *Urochloa humidicola* (Rendle) Morrone e Zuloaga, por Zuloaga e Morrone (2003), e *Brachiaria humidicola* (Rendle) Schweick é considerada sinonímia para a espécie.

Assim como outras espécies de *Urochloa* introduzidas no Brasil, *U. humidicola* faz parte do Programa de Melhoramento de Forrageiras Tropicais da

Embrapa Gado de Corte, em Campo Grande, Mato Grosso do Sul. O programa já realizou estudos de avaliações agronômicas e morfológicas, assim como estudos acerca do modo reprodutivo e da citogenética da espécie, e os resultados encontrados permitiram que o germoplasma desta forrageira fosse explorado e utilizado no melhoramento genético (JUNGMANN *et al.*, 2010).

Apesar da importância forrageira da espécie e do gênero *Urochloa*, poucos trabalhos de informações genéticas de *Urochloa humidicola* são relatados na literatura (SALGADO *et al.*, 2017). Geralmente, os estudos envolvem a avaliação das espécies do gênero, seja por meio de identificação da diversidade genética dentro do grupo (TRIVIÑO *et al.*, 2017; KUWI *et al.*, 2018), seja por validação de marcadores moleculares (SOUZA *et al.*, 2018). Alguns trabalhos específicos de sequenciamento do transcriptoma de *U. humidicola* foram relatados, como o de Vigna *et al.* (2016a), que avaliaram genes expressos em dois genótipos divergentes da espécie, um sexual e um apomítico facultativo, e foram identificados unigenes relacionados a biossíntese de lignina, a fotossíntese de C4 e a resposta ao estresse por encharcamento.

2.8 CARACTERIZAÇÃO DO GÊNERO *Cenchrus* E DA ESPÉCIE *Cenchrus purpureus*

O gênero *Cenchrus* pertence à família Poaceae, subfamília Panicoideae e tribo Paniceae (PEREIRA *et al.*, 2016a) e, nas últimas décadas, algumas alterações foram sugeridas ao grupo. Desde 2009, alguns estudos buscaram elucidar a relação entre *Cenchrus* e *Pennisetum*, dois gêneros considerados como intimamente relacionados (VERLOOVE, 2012; VELDKAMP, 2014). Dentre eles, destaca-se o trabalho realizado por Chemisquy *et al.* (2010), no qual avaliaram 51 espécies pertencentes a oito gêneros da subfamília Panicoideae: *Cenchrus*, *Pennisetum*, *Odontelytrum*, *Ixophorus*, *Paspalidium*, *Setaria*, *Stenotaphrum* e *Rupichloa*. Os resultados obtidos foram que os gêneros *Cenchrus* e *Pennisetum* formariam um grupo monofilético, os quais sustentaram fortemente a unificação dos dois grupos em um único gênero. Fundamentados nestes estudos, taxonomistas propuseram a existência apenas do gênero *Cenchrus* (CHEMISQUY *et al.* 2010; GUTIÉRREZ, 2015; PEREIRA

et al., 2016a), e uma nova nomenclatura para as espécies antes pertencentes ao gênero *Pennisetum* (PEREIRA *et al.*, 2016a).

Dadas estas mudanças, o capim-elefante (Figura 5) é descrito atualmente como *Cenchrus purpureus* (Schumach.) Morrone, sinonímia de *Pennisetum purpureum* Schumach. (STAPT E HUBBARD, 1934; PEREIRA *et al.*, 2016a). A espécie ocorre naturalmente em uma extensa área da África Oriental, principalmente nos territórios de Guiné, Moçambique, Angola, Zimbábue e sul do Quênia, sendo esta região considerada como o centro de origem e de diversidade genética do capim-elefante (FERREIRA E PEREIRA, 2005; CAVALCANTE E LIRA, 2010). É uma gramínea perene, de hábito e crescimento cespitoso, atingindo de dois a seis metros de altura (NASCIMENTO JUNIOR, 1981; ALCÂNTARA E BUFARAH, 1983; DERESZ, 1999b; PEREIRA *et al.*, 2016a). A espécie é predominantemente alógama, com alto nível de heterozigose. A propagação vegetativa é a forma mais comum de propagação, pois as sementes são geralmente de baixa germinação e, quando germinam, nascem plantas com baixo vigor (PEREIRA *et al.*, 2010). *C. purpureus* é alotetraploide, com $2n=4x=28$ (BRUNKEN, 1997; DAVIDE *et al.*, 2007), de número básico cromossômico $x=7$ (MANARA, 1973) e apresenta um dos seus genomas parcialmente homólogo ao do milho (*Cenchrus americanus*, $2n=2x=14$) (DUJARDIN E HANNA, 1985; REIS, 2016).

O capim-elefante é uma espécie com alta taxa fotossintética e eficiência no uso da água, pela característica das folhas estreitas e eretas que permitem maior penetração da luz através do perfilho vegetal (ALCÂNTARA E BUFARAH, 1983; ANDERSON *et al.*, 2008; KANDEL *et al.*, 2016). Essas características, associadas a boa qualidade, palatabilidade, vigor, persistência, boa adaptação aos diversos tipos de solo, e elevada produção de matéria seca (SOUZA SOBRINHO *et al.*, 2005), permitiram que a forrageira tivesse grande aceitação na alimentação bovina (HANNA E SOLLENBERGER, 2007; YASUDA *et al.*, 2014), contribuindo para o aumento da produção de leite (DERESZ, 1999a) e de carne no Brasil (PEREIRA *et al.*, 2001; SOUZA SOBRINHO *et al.*, 2005). Além disso, a forrageira também é utilizada na produção de bioenergia, pois possui alta eficiência na conversão de luz em biomassa, alta eficiência no uso da água e no uso de nitrogênio (TAYLOR *et al.*, 2010).



Figura 5: Representação da espécie *Cenchrus purpureus*.

Dada a importância dessa espécie em dois setores da economia, a Embrapa Gado de Leite desenvolve, desde 1991, um programa de melhoramento genético do capim-elefante (PEREIRA *et al.*, 2010). O melhoramento da espécie busca obter novas combinações gênicas para atender à demanda nacional por cultivares de forrageiras superiores, agregando características como maior velocidade de crescimento, tolerância a solos de baixa fertilidade, distribuição mais equitativa da produção de matéria seca durante o ano, além de resistência a cigarrinhas das pastagens (PEREIRA *et al.*, 2003; PEREIRA *et al.*, 2010; REIS, 2016). Na produção de energia mecânica, os materiais são avaliados observando-se a concentração de celulose, hemicelulose e lignina, que influenciam diretamente na eficiência da produção de etanol celulósico. No âmbito da produção de energia térmica, o melhoramento do capim-elefante visa materiais que apresentam alta relação de carbono:nitrogênio, elevado potencial de produção e baixo teor de proteína. Mais especificamente, os materiais são selecionados a partir de fatores como umidade, cinzas, matéria volátil, poder calorífico superior e poder calorífico inferior, além de

concentração de hidrogênio, nitrogênio, carbono e oxigênio, para que sejam utilizados como fonte de produção de geração de energia térmica (PRONOBIS, 2006).

Apesar de sua importância em diferentes setores da economia, principalmente no Brasil, o capim-elefante é considerado uma espécie órfã, pois são poucas informações encontradas na literatura a nível molecular. Os primeiros estudos de RNA-Seq envolvendo a espécie foram focados principalmente na identificação de marcadores moleculares. Zhou *et al.* (2018) sequenciaram o RNA de dois genótipos dessa espécie para seleção de marcadores moleculares específicos para capim-elefante. Nos últimos anos, novos estudos foram publicados voltados para a descoberta de genes envolvidos em rotas importantes. Zhang *et al.* (2020) estudaram diferentes tempos de desenvolvimentos em caules de capim-elefante para a compreensão de genes reguladores da síntese de moléculas da lignocelulose. Já Yan *et al.* (2021a) avaliaram duas cultivares contrastantes de *C. purpureus* para a identificação de genes associados ao processo de alongamento do internó, para o desenvolvimento de gramíneas de maior eficiência energética. Informações acerca do genoma sequenciado da espécie são ainda mais escassas, com o trabalho recentemente publicado por Yan *et al.* (2021b) o primeiro descrito na literatura. Por todos os motivos apresentados, é imprescindível que novas pesquisas tragam informações moleculares de capim-elefante, que serão fonte de estudos no melhoramento da espécie no desenvolvimento de cultivares.

2.9 MARCADORES MOLECULARES

Com a constante demanda de culturas que atendem a diferentes necessidades do mercado, melhoristas do mundo todo vêm, a cada ano, utilizando diferentes metodologias para a seleção dos melhores materiais. Características como estrutura da semente, cor da flor e hábito de crescimento podem ser usados como marcadores morfológicos de diferentes culturas a partir da visualização; porém, é dependente do estágio de crescimento vegetal e de fatores ambientais (NADEEM *et al.*, 2018). Marcadores citológicos, que avaliam variações presente no número, tamanho, formato e posição dos cromossomos, e marcadores bioquímicos, que estão

relacionados a isoenzimas (NADEEM *et al.*, 2018), são técnicas laboratoriais que auxiliaram na identificação de variedades genéticas de culturas como videira (ROYO *et al.*, 1997; IQBAL *et al.*, 2021), mandioca (SUMARANI *et al.*, 2004; IQBAL *et al.*, 2021) e alho (IPEK *et al.*, 2003; IQBAL *et al.*, 2021).

Nas últimas décadas, os marcadores moleculares são as principais ferramentas utilizadas por melhoristas a fim de se identificar características de interesse agrônomo de diversas espécies, como arroz (MACKILL *et al.*, 1999; GARRIDO-CARDENAS *et al.*, 2018), milho (ORTIZ, 2010; SUWARNO *et al.*, 2015; GARRIDO-CARDENAS *et al.*, 2018) e trigo (LANDJEVA *et al.*, 2007; GARRIDO-CARDENAS *et al.*, 2018). Estes marcadores são formados a partir de variações na sequência de nucleotídeos, e por isso permitem a investigação de polimorfismo presente nas sequências de diversos indivíduos (NADEEM *et al.*, 2018). Nesta categoria de marcadores, estão presentes técnicas que são baseadas na hibridização (LANDER E BOTSTEIN, 1989; GARRIDO-CARDENAS *et al.*, 2018), como *Restriction Fragment Length Polymorphism* (RFLP); técnicas baseadas em PCR (O'HANLON *et al.*, 2000; GARRIDO-CARDENAS *et al.*, 2018), como *Random Amplified Polymorphic DNA* (RAPD), *Amplified Fragment Length Polymorphism* (AFLP) e *Simple Sequence Repeat* (SSR); e técnica de marcadores baseados na sequência (GANAL *et al.*, 2009; GARRIDO-CARDENAS *et al.*, 2018), por exemplo *Single Nucleotide Polymorphism* (SNP).

SSR, ou microssatélites, são marcadores amplamente distribuídos no genoma, sendo também encontrados em cloroplastos (PROVAN *et al.*, 2001; GARRIDO-CARDENAS *et al.*, 2018), mitocôndrias (RAJENDRAKUMAR *et al.*, 2007; GARRIDO-CARDENAS *et al.*, 2018) e em genes codificantes de proteínas e *tags* de sequências expressas (MORGANTE *et al.*, 2002; NADEEM *et al.*, 2018). Por estar presente em diferentes organelas e moléculas celulares, somada a características como alto nível de polimorfismo e natureza co-dominante (ADAMS *et al.*, 2006; RODRIGUEZ *et al.*, 2010; WANG *et al.*, 2012; HSU *et al.*, 2013; ZHANG *et al.*, 2014; LIU *et al.*, 2020b), este marcador se tornou a principal ferramenta em diversos trabalhos. Na literatura, encontram-se trabalhos com SSR para estudos de filogenia e evolução de marcadores cloroplastidiais (WANG *et al.*, 2011; DONG *et al.*, 2012; NADEEM *et al.*, 2018), investigação de diversidade genética e heterose (WU *et al.*, 2013; NADEEM *et al.*, 2018), fluxo gênico (LIU *et al.*, 2020b; LU *et al.*, 2021), genética

de populações (ZHAI *et al.*, 2018; BELALIA *et al.*, 2019; LIU *et al.*, 2020b), e genética da conservação (YASODHA *et al.*, 2018; LIU *et al.*, 2020b). No campo do melhoramento vegetal, o marcador está presente em estudos de mapeamento de QTL (*Quantitative Trait Loci*) (NADEEM *et al.*, 2018; SU *et al.*, 2018; CHOI *et al.*, 2019) e no *fingerprint* do DNA, que se tornou importante na identificação de cultivares.

2.10 FINGERPRINT DO DNA

Diversas espécies vegetais de interesse econômico são objetos de estudo para o desenvolvimento de cultivares que apresentem características específicas, como por exemplo maior produção de matéria seca e resistência a estresse. A partir do desenvolvimento dessas cultivares, sementes e/ou mudas das cultivares são comercializadas. Porém, nem sempre o que é vendido apresenta as características iniciais das cultivares melhoradas.

Para garantir a pureza genética das sementes e mudas, e conseqüentemente proteger a propriedade intelectual, é necessário que seja feita uma identificação precisa da cultivar comercializada (SATTURU *et al.*, 2018; NIE *et al.*, 2019).

Inicialmente, a avaliação da qualidade e pureza de cultivares era baseada na observação de características morfológicas (HWANG *et al.*, 2020; IQBAL *et al.*, 2021), como formato, cor e tamanho de frutos, pubescência das folhas e número de flores por inflorescência (WALSH E HOOT, 2001; IQBAL *et al.*, 2021). Porém, essa classificação baseada na morfologia é um processo demorado, pois as características são observadas ao longo do crescimento vegetal (SATTURU *et al.*, 2018). Além disso, esses traços podem sofrer efeito da interação genótipo ambiente, dificultando ainda mais a identificação (BORA *et al.*, 2016; SATTURU *et al.*, 2018; IQBAL *et al.*, 2021).

A partir da necessidade de se classificar as cultivares de maneira mais precisa, eficiente e confiável (IQBAL *et al.*, 2019; IQBAL *et al.*, 2021; TIAN *et al.*, 2021), pesquisadores voltaram suas pesquisas na identificação de características únicas presentes no genótipo das plantas. Para isso, foram utilizados como ferramentas marcadores moleculares encontrados no DNA para a identificação de variedades

(JAMIL *et al.*, 2020a; IQBAL *et al.*, 2021) que ficou conhecido como *fingerprint* do DNA (IQBAL *et al.*, 2021). Marcadores moleculares como RAPD e AFLP foram, no passado, usados para a identificação de *fingerprint* no DNA de culturas importantes como milho, trigo, arroz, algodão e soja (SALEM *et al.*, 2007; ZARGAR *et al.*, 2017; IQBAL *et al.*, 2019; IQBAL *et al.*, 2021). Nos últimos anos, o *fingerprint* no DNA de cultivares é baseado na utilização de marcadores SSR. Algumas características como alto nível polimórfico, co-dominância e serem amplamente distribuídos no genoma (LIU *et al.*, 2020a; HWANG *et al.*, 2020) tornaram esse marcador confiável para a caracterização de cultivares (BANG *et al.*, 2011; HWANG *et al.*, 2020), como cevada, soja, trigo, palmeira, arroz, milho (WANG *et al.*, 2011; JAMIL *et al.*, 2020a; JAMIL *et al.*, 2020b; IQBAL *et al.*, 2021), hortaliças (TSAI *et al.*, 2013; XIE *et al.*, 2011a; XIE *et al.*, 2011b; NIE *et al.*, 2019) e forrageiras (TEHRANI *et al.*, 2009; HERMANN *et al.*, 2010; JIANG *et al.*, 2013; NIE *et al.*, 2019).

Uma barreira importante na descoberta de marcadores microssatélites em espécies vegetais é a identificação dessas regiões, que são realizadas geralmente nos genomas. O genoma sequenciado em banco de dados já é realidade em espécies importantes, como *Arabidopsis thaliana*, milho, arroz, trigo, soja, e muitos outros. Porém, em espécies pouco estudadas, o sequenciamento do genoma ainda não está disponível, pois necessita um maior investimento se comparado a outras técnicas de descoberta de marcadores. Como forma de contornar esse problema, o sequenciamento do RNA pode ser uma ferramenta mais rápida, fácil e de melhor custo-benefício (TAHERI *et al.*, 2018; LIU *et al.*, 2020a). Um estudo publicado por Liu *et al.* (2020a) desenvolveram 16 marcadores microssatélites a partir do sequenciamento do transcriptoma de *Angelica dahurica*, que foram importantes na investigação da diversidade genética da espécie e suas cultivares. Para a descoberta de SSRs em *Lolium multiflorum*, Pan *et al.* (2018) basearam-se no RNA-Seq para a identificação de 50 marcadores, que foram posteriormente testados em espécies de *Lolium* e *Festuca*.

Urochloa humidicola e *Cenchrus purpureus* são espécies utilizadas na alimentação de bovinos no Brasil, e que ao longo dos anos, foram desenvolvidas cultivares a fim de atender a necessidade do mercado de forrageiras. Segundo Vigna *et al.* (2016a), três cultivares de *Urochloa humidicola* estão disponíveis no Brasil (Tully, Llanero e BRS Tupi). Já para Capim-Elefante, quatro cultivares estão disponíveis

atualmente (BRS Kurumi, BRS Canará, BRS Capiacu e Pioneiro). Assim como visto em outras culturas, um dos problemas encontrados neste mercado é a garantia da pureza das cultivares, e um diagnóstico genético a partir da identificação do *fingerprint* do DNA exclusivo de cada cultivar pode ser uma ferramenta muito útil. Sendo assim, o RNA sequenciado das espécies para a investigação de genes relacionados a produção de lignina podem servir como base para a descoberta e o desenvolvimento de marcadores SSRs que, posteriormente, serão úteis na classificação das diversas cultivares.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Identificar genes relacionados a produção de lignina diferencialmente expressos e marcadores microssatélites nas espécies *Urochloa humidicola* e *Cenchrus purpureus*.

3.2 OBJETIVOS ESPECÍFICOS

1. Identificar genes diferencialmente expressos em amostras contrastantes para produção de lignina em *Urochloa humidicola* e *Cenchrus purpureus*;
2. Anotar as funções gênicas dos genes diferencialmente expressos identificados em *U. humidicola* e *C. purpureus*;
3. Identificar marcadores microssatélites nos transcriptomas de *U. humidicola* e *C. purpureus*;
4. Desenhar *primers* de microssatélites para amplificação dos marcadores moleculares em ambas as espécies;
5. Testar a amplificação dos pares de *primers* de microssatélites selecionados em *U. humidicola* e capim-elefante;
6. Avaliar o polimorfismo dos marcadores microssatélites identificados entre as cultivares de capim-elefante lançadas pela Embrapa.

CAPÍTULO 1: ANÁLISE DA EXPRESSÃO GÊNICA EM *Urochloa humidicola*

1.1 INTRODUÇÃO

O gênero *Urochloa* possui espécies importantes de gramíneas forrageiras tropicais na Austrália, leste da Ásia e na América do Sul (MILES *et al.*, 1996; KUWI *et al.*, 2018). No Brasil, as espécies do gênero contribuíram muito para o aumento da atividade pecuária em solos ácidos e de baixa fertilidade natural do Cerrado (DO VALLE *et al.*, 2009a), e estão presentes em 85% das pastagens cultivadas no país (DO VALLE *et al.*, 2008; VIGNA *et al.*, 2016a).

Dentre as espécies encontradas no Brasil, pode-se ressaltar *Urochloa humidicola*, que se adaptou a solos ácidos e áreas mal drenadas (KELLER-GREIN *et al.*, 1996; VIGNA *et al.*, 2016a) e a presença de amônia ou nitrato (CASTILLA E JACKSON, 1991; NUÑEZ *et al.*, 2018). É uma espécie preferencialmente apomítica, com ploidia variando de hexa a nanoploides (JUNGMANN *et al.*, 2009).

A espécie faz parte do Programa de Melhoramento de Forrageiras Tropicais da Embrapa Gado de Corte, que visa desenvolver cultivares melhoradas geneticamente que atendam às demandas do mercado de forrageiras (JUNGMANN *et al.*, 2010). Dentro do programa, algumas características são avaliadas, como a digestibilidade do material vegetal em rúmens de bovinos. Um material de melhor qualidade é aquele que possui maiores concentrações de compostos nutritivos, como proteínas e vitaminas, e menores concentrações de fibras, compostas principalmente por polímeros de lignina encontrados nas paredes celulares vegetais (WAGHORN E CLARK, 2004; LEE *et al.*, 2017). Assim, o conhecimento a partir da identificação de fatores que contribuem para a produção de lignina em *U. humidicola* pode ser uma importante ferramenta para o desenvolvimento de cultivares de melhor digestibilidade, com impactos diretos na produtividade da bovinocultura.

Ao longo dos anos, os processos envolvidos na produção de lignina foram revelados, e em espécies como *Populus* (YE *et al.*, 2020), sorgo (TETREULT *et al.*, 2020) e *Panicum virgatum* (RAO *et al.*, 2019), já existem estudos que revelam a expressão de genes já descritos como reguladores da via (BOERJAN *et al.*, 2003; XU *et al.*, 2009). A síntese de lignina é iniciada pela desaminação do aminoácido fenilalanina (KOUKOL E CONN, 1961; DA SILVA, 2019), catalisada pela enzima PAL (BOERJAN *et al.*, 2003). Em continuação ao processo, enzimas como 4CL, CCR, HCT, CAD, CCoAOMT, F5H e COMT estão presentes para a formação das unidades

p-hidroxifenil (H), siringil (S) e guaiacil (G). As unidades sintetizadas são então adicionadas ao polímero de lignina em construção, processo conhecido como lignificação (BOERJAN *et al.*, 2003).

A técnica de sequenciamento de RNA (RNA-seq) é frequentemente utilizada em trabalhos de análise de genes diferencialmente expressos (STARK *et al.*, 2019). A eficiência da técnica pode ser vista em estudos de espécies como arroz (KONG *et al.*, 2020), *Arabidopsis thaliana* (FARMER *et al.*, 2021; PARK *et al.*, 2021) e milho (FAN *et al.*, 2020), contribuindo no conhecimento de vias e expressão de genes de interesse para o desenvolvimento de variedades de espécies de interesse econômico.

O objetivo deste trabalho é conhecer o nível de expressão de genes associados a via de biosíntese de lignina, por meio do sequenciamento dos RNAs de oito genótipos selecionados de *Urochloa humidicola*, que apresentem diferenças na concentração de lignina.

1.2 MATERIAL E MÉTODOS

1.2.1 MATERIAL VEGETAL E QUANTIFICAÇÃO DE LIGNINA

A seleção dos genótipos de *U. humidicola* foi feita a partir de dados do histórico de produção de lignina (Tabela 1) realizado pelo Programa de Melhoramento de Forrageiras da Embrapa Gado de Corte, situada na cidade de Campo Grande, Mato Grosso do Sul. Anualmente, todos os materiais pertencentes ao programa de melhoramento são avaliados a campo com relação às características bromatológicas na estação das águas e de seca, 60 dias após o corte de uniformização. A quantificação de lignina é realizada por meio das metodologias de espectroscopia no infravermelho próximo (NIRS) e por via úmida (GOERING E VAN SOEST, 1967). Foram selecionados oito genótipos apomíticos, separados em dois grupos: grupo com maior produção de lignina (quatro genótipos) e grupo com menor produção de lignina (quatro genótipos).

Tabela 1: Lista dos oito genótipos de *Urochloa humidicola* pertencentes ao Programa de Melhoramento de Forrageiras da Embrapa Gado de Corte utilizados no sequenciamento de RNA.

Genótipo <i>U. humidicola</i>	Ploidia	Classificação - Produção de lignina
BRA000540	9X	Baixa
BRA005029	6X	Baixa
BRA005053	9X	Alta
BRA004959	9X	Baixa
BRA001929	9X	Alta
6 A – H048	7X	Alta
BRA005100	6X	Baixa
BRA006149	9X	Alta

1.2.2 COLETA DE AMOSTRAS E EXTRAÇÃO DE RNA

A coleta para a realização da extração do RNA total foi realizada a partir da segunda folha completamente expandida (Figura 4), e o material utilizado foi previamente limpo com álcool 70% e RNase Away® (Life Technologies, Carlsbad, Estados Unidos). Em seguida, as folhas selecionadas foram colocadas em tubos de 2mL com 1mL de solução RNA/later® (Life Technologies, Carlsbad, Estados Unidos), e armazenadas a -20°C.

As folhas de todos os acessos foram então maceradas em nitrogênio líquido, sendo todo o material utilizado na maceração previamente higienizado com o reagente RNase Away® (Life Technologies, Carlsbad, Estados Unidos), para eliminar qualquer contaminação da enzima RNase no processo.

O RNA das folhas jovens de *U. humidicola* maceradas foi extraído utilizando RNeasy® Plant Mini Kit (Qiagen, Hilden, Alemanha), seguindo o protocolo do kit. Ao final do processo, o RNA extraído foi ressuspendido em 50µL de água livre de RNase fornecido pelo kit, e aliqotado em cinco tubos de 0,5mL, com os seguintes volumes: um tubo com 3µL, um tubo com 2µL, e três tubos com 15µL cada.

A concentração de RNA foi estimada por leitura em Nanodrop 1000 (Thermo Fisher, Waltham, Estados Unidos), com os comprimentos de onda de 260 e 280 nm.

A integridade do RNA foi estimada utilizando o equipamento Agilent 2100 Bioanalyzer e kit RNA 600 Nano (Agilent Technologies, Santa Clara, Estados Unidos), de acordo com especificações do fabricante.



Figura 6: Morfologia de um perfilho em gramíneas. Adaptado de Matthew *et al.* (2001).

Para a avaliação das quantidades de FDA e lignina no momento da coleta, as folhas selecionadas foram avaliadas seguindo a metodologia proposta por Goering e Van Soest (1967).

1.2.3 MONTAGEM DE BIBLIOTECA DE cDNA

As amostras selecionadas para o sequenciamento foram baseadas nos valores de *RIN* superiores a 6,4 e quantidade de RNA superior a 1000 ng. Seguindo esse critério, as oito amostras de *U. humidicola* foram sequenciadas.

As etapas de conversão de RNA em cDNA e construção das bibliotecas e sequenciamento utilizando a tecnologia Illumina HiSeq2500 foram desenvolvidas pela empresa BIODOME.

As bibliotecas de DNA complementar (cDNA) foram preparadas de acordo com o protocolo Illumina TruSeq Stranded mRNA Sample Prep LT. Inicialmente, as

amostras de RNA foram purificadas usando esferas magnéticas ligadas a oligo-dT, que tem como função a separação do RNA mensageiro (mRNA) pela ligação a cauda poli-A contida na molécula. Em seguida, o mRNA purificado foi fragmentado e ligado a *primers* para seguir com a síntese de cDNA. A primeira fita de cDNA foi sintetizada a partir dos RNAs fragmentados, utilizando a enzima *SuperScript II Reverse Transcriptase* (Illumina). A síntese da segunda fita de cDNA foi realizada por meio do *Second Strand Marking Master Mix*, que ao final desta etapa, produziu cDNA com extremidades abruptas ligadas a esferas magnéticas. O cDNA foi então adenilado nas extremidades 3' das fitas por meio do *A-Tailing Mix* (Illumina), evitando assim que as fitas se ligassem durante a reação de ligação de adaptadores. Um nucleotídeo de timina correspondente na porção 3' dos adaptadores forneceu a ligação complementar aos fragmentos adenilados, evitando a formação de quimeras no processo. Com o cDNA ligado aos adaptadores, foi feita a hibridização dos fragmentos de cDNA na *flow cell*, realizada por meio de *Ligation Mix* (Illumina), que permite que os adaptadores encontrados nos fragmentos de cDNA se hibridizem aos adaptadores encontrados na *flow cell*. Ao final, as bibliotecas foram purificadas e amplificadas por PCR, pela ligação de *primers* aos adaptadores, na seguinte condição: 98°C por 30 s, 15 ciclos de 98°C por 10 s, 60°C por 30 s e 72°C por 30 s, seguido por 72°C por 5 min.

A biblioteca produzida foi avaliada no equipamento *Bioanalyzer 2100* (Agilent Technologies) utilizando DNA 1000 chip (Agilent Technologies), para avaliação da qualidade da biblioteca. As amostras consideradas ótimas apresentaram uma banda de aproximadamente 260 pb.

1.2.4 SEQUENCIAMENTO DE mRNA (RNA-Seq)

O sequenciamento foi realizado com tecnologia de nova geração, utilizando o equipamento *HiSeq 2500* (Illumina). A metodologia baseou-se no sequenciamento por síntese, no qual cada base nitrogenada é identificada logo que é adicionada à cadeia recém-formada, por meio da enzima DNA polimerase. As amostras ligadas a adaptadores específicos são complementares a oligos presentes na superfície da *HiSeq Flow Cell v4* (Illumina) e amplificadas, produzindo *clusters* (agrupamentos) de

fragmentos iguais. Em seguida, a sequência dos *clusters* foi identificada por meio do *HiSeq SBS Kit v4* (Illumina), com leituras pareadas de 100 pb (2x), de acordo com recomendações do fabricante. Essa metodologia permitiu a detecção dos nucleotídeos, complementares às sequências dos *clusters* e marcados por fluorescência, após sua incorporação a cada ciclo da reação, pela excitação da fluorescência acoplada.

1.2.5 ANÁLISE DE DADOS

1.2.5.1 Controle de Qualidade das bibliotecas

Após o sequenciamento, as bibliotecas geradas foram então avaliadas quanto à qualidade do sequenciamento. O controle de qualidade foi realizado por meio do *software* FastQC v0.11.8 (Babraham Bioinformatics, Babraham, Inglaterra), utilizando os arquivos de formato FASTQ. Para cada amostra analisada, um gráfico de Whisker foi gerado, utilizando o *Phred quality score*. O valor de *Phred* avalia a confiabilidade dos dados, atribuindo um valor de probabilidade de erro do sequenciamento para cada nucleotídeo.

1.2.5.2 Montagem *de novo* do transcriptoma

U. humidicola é uma espécie que não apresenta genoma de referência relatado na literatura. Em casos similares, utiliza-se de recursos como a montagem *de novo*, ou seja, as sequências do RNA em estudo são usadas para a construção do transcriptoma.

Neste trabalho, foram utilizadas duas abordagens: a primeira, pelo *software* Trinity v 2.8.4 (GRABHERR *et al.*, 2011), produzindo o transcriptoma a partir da

construção e processamento de grafos de Bruijn. A esta estratégia foram adicionados um comando da ferramenta Trimmomatic (BOLGER *et al.*, 2014) (--quality_trimming_params "ILLUMINACLIP:/apps/trinity/20180912-2.8.4/share/trimmomatic/adapters/TruSeq3-PE.fa:3:20:10:1:true SLIDINGWINDOW:4:13 LEADING:3 TRAILING:3 MINLEN:50"), que elimina possíveis resquícios de adaptadores utilizados no sequenciamento do RNA, e uma etapa de normalização para cada par de .fastq analisado (--normalize_by_read_set).

A segunda abordagem foi realizada pelo *assembler* Stringtie v 2.0.4 (PERTEA *et al.*, 2015), que utiliza arquivos no formato BAM para a montagem. Este formato de arquivo é gerado em programas de alinhamento como o STAR v 2.5.2b (DOBIN *et al.*, 2013), que realiza sua função a partir da indexação de um genoma de referência. Devido a ausência de um genoma de referência, como foi relatado anteriormente, foi selecionada a espécie *Urochloa ruzizensis* (GCA_015476505.1) como referência para *Urochloa humidicola*, já que é considerada uma espécie próxima.

Um único transcriptoma foi formado para ser a base das próximas análises, combinando os dados produzidos pelo Trinity e Stringtie. Para essa etapa, a ferramenta PASA v 2.2 (HAAS *et al.*, 2003) foi adotada. Para informar que o transcriptoma pelo Trinity foi montada de forma direcionada, foi adicionado o comando --transcribed_is_aligned_orient, além de informar e classificar todas as variações de splicing alternativos, os quais não ficam totalmente evidentes nas montagens realizadas, por meio do comando --ALT_SPLICE.

1.2.5.3 Indexação e Mapeamento de sequências

O mapeamento das sequências foi realizado para fornecer os dados para a montagem *de novo* pelo programa Stringtie, como foi acima citado. Por meio do programa STAR v 2.5.2b (DOBIN *et al.*, 2013), as oito amostras de *U. humidicola* foram alinhadas no genoma de *U. ruzizensis*. Inicialmente, o genoma de referência foi indexado, e em seguida o alinhamento de cada amostra foi realizado. Em ambas as etapas, foram seguidos os comandos padrões do programa.

1.2.5.4 Quantificação de transcritos

As *reads* sequenciadas foram quantificadas para calcular a expressão dos transcritos. A ferramenta Salmon v 0.7.2 (PATRO *et al.*, 2017) foi utilizada, iniciada pela indexação do transcriptoma gerado pela ferramenta PASA. Após esta etapa, as *reads paired end* de cada amostra foram quantificadas, e os dados obtidos serviram como arquivos de entrada para a análise dos genes diferencialmente expressos. Em ambas as etapas, foram seguidos os comandos padrões da ferramenta.

1.2.5.5 Identificação de genes diferencialmente expressos

O programa R v3.6.1 (<http://www.R-project.org>), por meio do pacote “DESeq2” (LOVE *et al.*, 2014) foi utilizado na identificação dos genes diferencialmente expressos em *U. humidicola*. O pacote tem como funções realizar a normalização dos dados fornecidos, garantindo melhor acurácia das amostras analisadas, cálculo das expressões gênicas e realização de testes para avaliação das diferenças de expressão. Para corrigir possíveis erros dos testes, devido a análise de múltiplas comparações, uma etapa adicional de correção *False Discovery Rate* (FDR) é realizada. A partir desse conjunto de dados gerados, os genes que apresentaram $p \leq 0.05$ foram considerados significativamente expressos nas amostras.

1.2.5.6 Anotação dos genes

As funções dos genes encontrados no transcriptoma de *U. humidicola* foram identificadas utilizando o software Trinotate v 3.2.0 (BRYANT *et al.*, 2017), que combina informações dos bancos de dados da *Kyoto Encyclopedia of Genes and Genomes* (KEGG) (KANEHISA *et al.*, 2012) e NCBI/BLAST v 2.9.0 (ALTSCHUL *et al.*,

1990). Além de dados gênicos, o software identifica proteínas, utilizando os bancos de dados do TrEMBL/SwissProt (THE UNIPROT CONSORTIUM, 2021), Pfam (PUNTA *et al.*, 2012) e NCBI/BLAST (ALTSCHUL *et al.*, 1990), e regiões de transmembrana por meio do banco de dados do TmHMM (KROGH *et al.*, 2001). Ao final das análises, todos os dados são compilados em um único banco de dados SQLite (<https://www.sqlite.org>), e exportados para uma planilha do Excel (Microsoft Corporation, Albuquerque, Estados Unidos). Em todas as etapas de busca no banco de dados, foram seguidos os comandos padrões disponibilizados pelo software Trinotate.

1.2.5.7 Identificação dos genes diferencialmente expressos e suas anotações

As planilhas geradas pela análise dos genes diferencialmente expressos e pela anotação gênica foram combinadas, a fim de se identificar as funções dos genes considerados significativos. Essa combinação foi realizada pelo comando *merge* encontrada no programa R (<http://www.R-project.org>).

1.3 RESULTADOS E DISCUSSÃO

1.3.1 ANÁLISE BROMATOLÓGICA

Os valores das quantidades de FDA e lignina para os oito genótipos avaliados estão apresentados na Tabela 2. A quantidade de FDA variou de 339 g/kg MS (BRA005053) a 385 g/kg MS (6A – H48), e a quantidade de lignina variou de 22 g/kg MS (BRA005100) a 33 g/kg MS (BRA006149).

Tabela 2: Dados da avaliação bromatológica dos oito genótipos de *Urochloa humidicola*, com os valores de fibra em detergente ácido (FDA) e lignina, em g/kg MS, e classificação dos genótipos com relação ao teor de lignina.

Amostra	FDA (g/kg MS)	Lignina (g/kg MS)	Teor Lignina
BRA000540	359	26	Baixo
BRA005029	366	26	Baixo
BRA005053	339	29	Alto
BRA004959	344	24	Baixo
BRA001929	345	32	Alto
6A – H048	385	29	Alto
BRA005100	335	22	Baixo
BRA006149	380	33	Alto

As concentrações aqui encontradas de FDA e lignina são próximos ao mínimo encontrado na literatura com estudos envolvendo o gênero *Urochloa*. Guimarães (2019) avaliou 169 amostras de espécies do gênero *Urochloa* sob diferentes condições de manejo. As coletas foram realizadas em quatro unidades da Embrapa, localizadas nas cidades de Campo Grande – MS, Sete Lagoas – MG, Coronel Pacheco – MG e São Carlos – SP. Os valores mínimos de FDA e lignina em detergente ácido analisados neste estudo foram de 314,31 g/kg MS e 37,88 g/kg MS, respectivamente. O estudo conduzido por Monrroy *et al.* (2017) identificaram uma variação de FDA entre 285 e 455 g/kg MS, enquanto Caixeta *et al.* (2013) encontraram

valores de FDA na faixa de 263 a 522 g/kg MS. Na avaliação da quantidade de lignina nas cultivares Xaraés, Piatã e Marandu, pertencentes a espécie *Urochloa brizantha*, Andrade e Assis (2008) descreveram uma variação de 24,6 a 31 g/kg MS na análise realizada por 3 anos, entre seca e águas. Já Caixeta *et al.* (2013) avaliaram 147 amostras do gênero, e as análises de produção de lignina revelaram uma faixa de produção entre 16 e 96 g/kg MS.

Uma justificativa com relação ao valor mínimo encontrado das quantidades de FDA e lignina é o material avaliado. Em nosso trabalho, a análise bromatológica foi realizada a partir da coleta de folhas, 60 dias após o corte de uniformização. Nesta porção vegetal, o nível de lignificação é baixo se comparado ao nível encontrado no caule (RUEDA *et al.*, 2020). Por isso, quando se compara aos valores encontrados na literatura, que analisaram a planta inteira em diferentes estágios de desenvolvimento, é esperado que as quantidades de FDA e lignina sejam abaixo das médias encontradas, chegando próximo ao valor mínimo.

1.3.2 CONTROLE DE QUALIDADE DAS BIBLIOTECAS E MONTAGEM DE NOVO DO TRANSCRIPTOMA

Oito bibliotecas de RNA-Seq foram sequenciadas, correspondente as quatro amostras com alta produção de lignina e quatro com baixa produção de lignina. O sequenciamento gerou, ao final, 123 milhões de *reads paired-ends* (PE) (Tabela 3), de tamanho 100 pb, com média de 15 milhões de *reads* por biblioteca.

Tabela 3: Total de *reads* obtidas em *Urochloa humidicola* para cada biblioteca gerada, percentual de CG e total de *reads* geral.

Genótipo	Total de reads	GC (%)
BRA000540	15.884.843	52
BRA005029	15.331.128	51
BRA005053	13.150.706	51
BRA004959	14.838.208	51
BRA001929	14.470.399	51
6A - H48	15.904.731	51
BRA005100	17.130.066	50
BRA006149	17.241.909	52
Total	123.951.990	

A avaliação da qualidade do sequenciamento das amostras BRA000540 e BRA004959 pode ser vista na Figura 5. O resultado aqui exibido foi semelhante ao sequenciamento encontrado para as outras seis amostras estudadas. O eixo y representa o valor de qualidade, e quanto maior o valor encontrado, melhor é a confiança na *base call*. O valor da média da qualidade é representado pela linha central azul, e ficou na faixa de 32 a 36 para todas as amostras. Sendo assim, foi possível confirmar que o sequenciamento foi de boa qualidade, e as sequências obtidas foram usadas para as análises de dados.

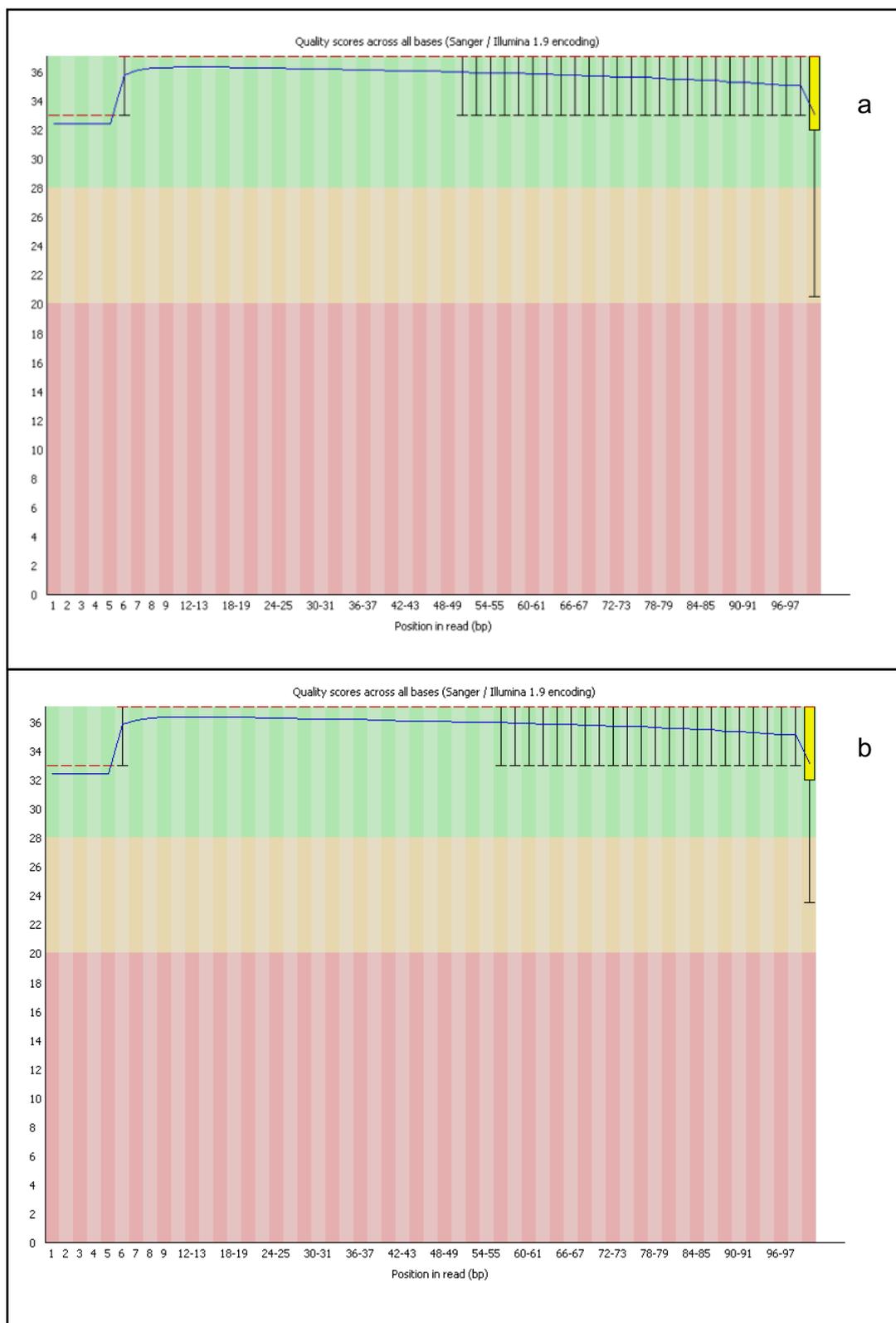


Figura 7: Gráfico de Whisker do sequenciamento do transcriptoma dos acessos de *Urochloa humidicola* BRA000540 (a) e BRA004959 (b). y é o eixo dos valores de qualidades. O fundo do gráfico divide o eixo em três níveis de confiança: muito confiável (verde), qualidade razoável (laranja) e baixa qualidade (vermelho). A linha central azul representa o valor da média da qualidade. A caixa amarela representa a amplitude interquartil (25-75%). Fonte: FastQC A Quality Control tool for High Throughput Sequence Data <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> por S. Andrews.

As duas abordagens de montagem *de novo* geraram quantidades de transcritos com valores discrepantes. Na abordagem Trinity, 641.069 transcritos foram montados, enquanto a abordagem Stringtie gerou 43.682 transcritos. Esta diferença nos valores é esperada, pois as abordagens realizam métodos de montagens diferentes. A montagem pelo Trinity é realizada sem ter alguma referência, e por isso o número de transcritos é superestimado ao que seria o número real. Isto permite que seja avaliado o máximo de possibilidades de transcritos possíveis da espécie analisada. Já a montagem Stringtie utiliza um genoma de referência para montar o transcriptoma. Consequentemente, os transcritos montados por essa abordagem são os mais próximos do transcriptoma real, e o número de transcritos é menor comparado a abordagem Trinity. Outro fator que influencia essa diferença está relacionado a ploidia dos genótipos avaliados de *U. humidicola* e o genoma de referência *U. ruziziensis*. A primeira espécie é classificada como hexaploide, enquanto a segunda é uma espécie diploide. Comparando-se o número de transcritos de indivíduos diploides e autotetraploides de espécies do gênero *Acer* e *Vaccinium*, Madritsch *et al.* (2021) relataram que, ao se utilizar Trinity para a montagem do transcriptoma, o número de transcritos de indivíduos autotetraploides quase duplica em relação aos diploides. Como a montagem pelo Stringtie é baseada no genoma diploide de *U. ruziziensis*, essas diferenças encontradas podem ser justificadas pelas diferentes ploidias encontradas. A escolha de se fazer duas montagens do transcriptoma de *Urochloa humidicola* é uma tentativa de avaliar o maior número possível de genes expressos da espécie no momento da coleta das folhas. Balmant *et al.* (2020) utilizou metodologia semelhante para avaliar o transcriptoma do xilema de uma população de *Populus deltoides*. As montagens dos transcriptomas foram realizados por meio das plataformas Cufflink, Stringtie e Trinity, para maximizar a detecção de novos genes.

Um transcriptoma final de *U. humidicola*, feito pela combinação das montagens do Trinity e Stringtie por meio do PASA, foi composto por 48.695 transcritos montados, um número mais próximo àquele encontrado na montagem pelo Stringtie. O número de transcritos encontrados aqui foi inferior ao encontrado por Vigna *et al.* (2016a), no qual 76.196 transcritos foram montados por meio do Trinity a partir de dois genótipos divergentes de *U. humidicola*. As análises em sequência para a identificação dos genes diferencialmente expressos foram realizadas a partir do transcriptoma final gerado pelo PASA.

1.3.3 INDEXAÇÃO E MAPEAMENTO DE SEQUÊNCIAS

O mapeamento das *reads* foi realizado a partir do genoma de *Urochloa ruzizensis*, espécie mais próxima que apresenta genoma sequenciado. Como pode ser visto na Tabela 4, as porcentagens de *reads* mapeadas variaram de 42,88% (BRA005053) a 46,73% (6A – H48), representando uma baixa porcentagem de alinhamento. Esse resultado é esperado por dois motivos principais: (i) alinhamentos realizados em um genoma de espécie diferente revelam somente regiões conservadas em ambas as espécies; e (ii) como foi relatado por Pessoa-Filho *et al.* (2017), que estudaram a filogenia de quatro espécies de *Urochloa* (*U. ruzizensis*, *U. decumbens*, *U. brizantha* e *U. humidicola*) a partir do sequenciamento de DNA cloroplastidial, *Urochloa humidicola* é considerada um táxon irmão de um clado fortemente sustentado formado por *U. ruzizensis*, *U. decumbens* e *U. brizantha*.

Tabela 4: Quantidade e porcentagem de *reads* sequenciadas de *U. humidicola* mapeadas utilizando como referência *U. ruzizensis*.

Biblioteca	Reads mapeadas (%)
BRA000540	7.086.650 (44,61%)
BRA005029	6.868.301 (44,8%)
BRA005053	5.823.034 (44,28%)
BRA004959	6.362.645 (42,88%)
BRA001929	6.374.367 (44,05%)
6A – H048	7.432.201 (46,73%)
BRA005100	7.633.055 (44,56%)
BRA006149	7.702.398(44,67%)

1.3.4 IDENTIFICAÇÃO DOS GENES DIFERENCIALMENTE EXPRESSOS E SUAS FUNÇÕES GÊNICAS

Para a análise dos genes diferencialmente expressos (DE), foram usados como contrastes a produção de lignina (alta produção x baixa produção). Os genótipos considerados de alto teor de lignina foram BRA005053, BRA001929, H048 e BRA006149; contrastando, os genótipos de baixo teor de lignina foram: BRA000540, BRA005029, BRA004959 e BRA005100. Dos 48.695 genes avaliados para expressão diferencial, 196 genes (0,4%) foram considerados significativos por apresentarem $p < 0,05$. Destes, 109 genes foram *up* regulados, e 87 genes foram *down* regulados nos genótipos de alta produção de lignina.

O sequenciamento de moléculas de DNA e RNA produz, ao final, sequências de funções gênicas desconhecidas. Para a identificação destes elementos funcionais no genoma, é utilizado o processo de anotação gênica, que vêm se aperfeiçoando nas últimas décadas (ABRIL E CASTELLANO, 2019). Neste trabalho, a anotação gênica realizada permitiu a identificação dos genes em diferentes bancos de dados. Os genes que foram anotados foram identificados principalmente nos bancos de dados do Pfam, KEGG e Gene Ontology (GO)/BLAST, por serem atualmente os bancos mais informativos acerca da anotação gênica. A partir deste critério, 39.820 genes foram anotados, incluindo suas isoformas.

Os dados da análise dos genes DE foram combinados aos encontrados na anotação gênica. Ao $p < 0,05$, 189 genes (0,4%) foram anotados para os bancos de dados do Pfam, KEGG e GO/BLAST.

Por meio do Pfam, foi possível anotar genes envolvidos em diversas funções celulares, mas em destaque os genes citocromo P450 (PF00067.22) e pectinacetilesterase (PF03283.13), ambos *down* regulados nos genótipos de alto teor de lignina. As enzimas do citocromo P450 possuem funções distintas nos organismos vegetais, incluindo sua função na via da lignina. Estudos em *Arabidopsis* revelaram que as principais enzimas P450 são ácidos cinâmico 4-hidroxilase (C4H), *p*-curamato 3-hidroxilase (C3H) e ferulato 5-hidroxilase (F5H) (GOU *et al.*, 2018; VANHOLME *et al.*, 2019). A enzima C4H está envolvida no início da via do fenilpropanoide, no

processo de produção de ácido p-cumárico, um precursor de compostos fenólicos como flavonoides e lignina (VOGT, 2010; GOU *et al.*, 2018). C3H e F5H são enzimas que estão diretamente envolvidas na produção dos monolignóis G e S, conhecidos como formadores do polímero de lignina. Apesar de ser conhecida a sua participação na via de lignina, somente com os estudos aqui realizados não é possível afirmar se os genes do citocromo P450 foram diferencialmente expressos devido ao contraste inicialmente estabelecido. Já a enzima pectina acetilsterase (PAE) está envolvida na modulação da acetilação da pectina, que são polissacarídeos estruturais na formação da parede celular primária e representam até um terço da massa seca de plantas (PHILIPPE *et al.*, 2017). Também pode-se destacar genes envolvidos na síntese de terpeno (PF01397.21 e PF03936.16), domínio de ligação de amido (PF00686.19) e transportador de açúcar (PF00083.24).

Pelo banco KEGG foram identificadas três enzimas envolvidas na formação da parede celular: pectina acetilsterase (*ath*:AT3G09410), 4-cumamato-CoA ligase (4CL) (*osa*:4331650) e xiloglucano endotransglicosilase (XET) (*ath*:AT4G25810). PAE, assim como descrito acima, está envolvida no processo de formação da parede celular, principalmente no crescimento vegetal e em resposta a estímulo ambiental (GOU *et al.*, 2018). A enzima 4CL participa de diferentes formas na via do fenilpropanoide, em resposta a estímulos de estresses bióticos e abióticos (LAVHALE *et al.*, 2018). Na via da biossíntese de lignina, a enzima está envolvida na produção de p-cumaroil-CoA, precursor do álcool p-cumaril, que é convertido na unidade H na parede celular secundária (CESARINO *et al.*, 2016; VANHOLME *et al.*, 2019). XET é uma enzima que está diretamente ligada na montagem e desmontagem da parede celular, com atividade catalítica no xiloglucano, um dos componentes da hemicelulose (MORALES-QUINTANA *et al.*, 2020). Esta estrutura pode ser encontrada tanto na parede celular primária quanto na secundária, em conjunto com a celulose e a lignina. Também pode-se destacar os genes anotados de alfa-terpineol sintase (*zma*:100125649), jasmonoil-L-aminoácido sintetase (*osa*:4339756) e isopreno sintase (*rcu*:8260996).

O banco de dados de GO/BLAST separa os genes em três categorias: componente celular, função molecular e processo biológico. Dentro do grupo processo biológico, alguns genes apresentam ligação a produção de lignina em plantas: formação do padrão do tecido vascular de caule (GO:0010222), organização

(GO:0071555) e espessamento (GO:0052386) da parede celular, processo metabólico de xiloglucano (GO:0010411) e processo de biossíntese de flavonoide (GO:0009813). Este último grupo gênico está ligado a produção do maior grupo de compostos fenólicos, e o mais diverso metabólito secundário encontrado em plantas (NABAVI *et al.*, 2018). A produção do flavonoide é iniciada pela via do fenilpropanoide, assim como a síntese de lignina (FRASER E CHAPPLE, 2011, GENG *et al.*, 2020). Por isso, alguns mecanismos foram observados a fim de regular a produção de cada composto. O trabalho de Geng *et al.* (2020) revelou que o fator de transcrição MYB4, que pertence a família de MYB, é responsável pela inibição da via do flavonoide e consequente otimização da via de lignina. Além dos genes relatados, foram identificados genes de respostas a sacarose (GO:0009744) e ácido jasmônico (GO:0009753), processos biossintéticos de terpenoide (GO:0016114), ácido jasmônico (GO:0009695) e brassinoesteroide (GO:0016132), processos metabólicos de esterol (GO:0016125) e ácido jasmônico (GO:0009694), e atividade de terpeno sintase (GO:0010333).

Os resultados aqui obtidos na análise de expressão gênica, somada a anotação gênica, foram informativos acerca de genes relacionados diretamente ou indiretamente a via do metabolismo de lignina. Na literatura, diversos estudos da análise de expressão gênica a partir do sequenciamento do RNA foram conduzidos a fim de se investigar e identificar genes importantes em vias metabólicas específicas. Yin *et al.* (2021) avaliaram o transcriptoma da cultivar *japonica* de arroz (*Oriza sativa*) para identificar genes responsáveis pela regulação da floração. Já Qiu *et al.* (2020) estudaram *Fetusca ovina* a fim de se investigar genes ligados a estresse hídrico associado a condições semelhantes a seca. A elucidação da expressão gênica em espécies de interesse econômico pode servir como base de estudos dentro de programas de melhoramento, como a manipulação gênica, que estão em constante evolução, e a identificação de marcadores específicos. Em *Urochloa humidicola*, Vigna *et al.* (2016a) sequenciaram o RNA da cultivar Koronivia, como forma de entender os genes expressos, relacionar a características de interesse, e utilizar como marcadores específicos dessa cultivar que não era disponível no Brasil. Os genes diferencialmente expressos nas oito amostras avaliadas, e que estão relacionadas na via de metabolismo de lignina, poderão ser importantes ferramentas em estudos de manipulação genética no desenvolvimento de cultivares de *U. humidicola*, nas quais

reunirão características como melhor eficiência na digestibilidade e menores concentrações de FDA e lignina.

1.4 CONCLUSÃO

A análise de genes DE a partir do sequenciamento do RNA de oito genótipos de *Urochloa humidicola* revelou genes identificados por interferir na via de biossíntese de lignina. Em destaque, genes que expressam a enzima 4CL e aqueles envolvidos na biossíntese de flavonoides, pois são importantes na regulação da produção de lignina em vegetais. Estes genes podem ser genes chave nas estratégias de melhoramento da espécie, no desenvolvimento de cultivares com maior eficiência na digestibilidade.

CAPÍTULO 2: ANÁLISE DA EXPRESSÃO GÊNICA EM *Cenchrus purpureus*

2.1 INTRODUÇÃO

Capim-Elefante (*Cenchrus purpureus*) é uma espécie predominantemente alógama, que se reproduz preferencialmente por propagação vegetativa (PEREIRA *et al.*, 2010). Tem o centro de origem e de diversidade genética em uma extensa área da África Oriental, em territórios de Guiné, Moçambique, Angola, Zimbábue e sul do Quênia (FERREIRA E PEREIRA, 2005; CAVALCANTE E LIRA, 2010). A espécie reúne características como alta taxa fotossintética, boa qualidade, palatabilidade e elevada produção de matéria seca (SOUZA SOBRINHO *et al.*, 2005), o que a tornou importante como fonte de alimento para a população bovina no Brasil (HANNA E SOLLENBERGER, 2007; YASUDA *et al.*, 2014) e produção de bioenergia (TAYLOR *et al.*, 2010).

No processo de alimentação em ruminantes, a avaliação da digestibilidade do alimento é importante para se obter uma maior eficiência nutricional (REIS *et al.*, 2016; WEISS E HALL, 2020). Nesta avaliação, a concentração de lignina é uma das principais características investigadas, já que a molécula presente na parede celular de vegetais faz parte da porção fibrosa do alimento, que é conhecida por ser de difícil digestão. (WAGHORN E CLARK, 2004; LEE *et al.*, 2017). Na produção de bioenergia, maiores concentrações de lignina contribuem para a recalcitrância da biomassa, servindo como barreira física ao acesso de enzimas envolvidas no processo de conversão do material vegetal para etanol de segunda geração (2G) (ZENG *et al.*, 2020). Em contrapartida, uma biomassa com alta concentração de lignina contribui para o aumento do poder calorífico quando o material é usado na combustão para a produção de energia térmica (FERNANDES *et al.*, 2020).

Com o intuito de desenvolver cultivares que atendam a demanda do mercado, a Embrapa Gado de Leite possui um Programa de Melhoramento da espécie que, desde 1991, foi responsável pelo lançamento de diversas cultivares melhoradas. Para melhor atender ao mercado de alimentação animal e de bioenergia, é fundamental um estudo detalhado sobre a via de biossíntese de lignina. Os resultados obtidos contribuirão para o desenvolvimento de cultivares melhoradas a partir da manipulação de genes importantes na via de biossíntese dessa molécula.

Enzimas envolvidas no processo de síntese de lignina já foram descritas e bem caracterizadas (BOERJAN *et al.*, 2003; XU *et al.*, 2009). Inicialmente, o aminoácido fenilalanina sofre uma desaminação pela enzima PAL, e ao final do processo da via, são produzidas as unidades *p*-hidroxifenil (H), siringil (S) e guaiacil (G). Já na parede celular secundária, as enzimas lacase e peroxidase atuam para a formação do polímero de lignina (BOERJAN *et al.*, 2003). Em capim-elefante, Zhang *et al.* (2020) avaliaram diferentes estágios de desenvolvimento de caules da espécie para identificação da expressão de genes relacionados a síntese de lignocelulose. Como resultado, os pesquisadores identificaram expressões de genes ligados ao metabolismo de fenilpropanoide, amido e sacarose, os quais são relacionados ao desenvolvimento da parede celular, com a síntese de hemicelulose, celulose e lignina. Na síntese de lignina, foi possível identificar alta expressão dos genes CAD, 4CL, C4H, PAL, CCR, FSH e CCoAOMT em caules maduros de capim-elefante.

O objetivo deste estudo é identificar o nível de expressão de genes importantes na biossíntese de lignina, a partir do sequenciamento do RNA de acessos de capim-elefante caracterizados por alta e baixa produção de lignina. Os resultados obtidos serão fundamentais para o desenvolvimento de cultivares com concentrações ideais de lignina.

2.2 MATERIAL E MÉTODOS

2.2.1 MATERIAL VEGETAL

Inicialmente, foram selecionados 20 acessos pertencentes ao Banco Ativo de Germoplasma de Capim-Elefante da Embrapa Gado de Leite (BAGCE) a partir do histórico da produção de lignina avaliada durante seis anos, nas estações secas e chuvosas (ROCHA *et al.*, 2019) (Tabela 5).

Todos os acessos do BAGCE foram caracterizados no primeiro ano de avaliação bromatológica seguindo a metodologia proposta por Goering e Van Soest (1967) para determinação das concentrações de FDA e lignina. Nos demais cinco anos de avaliação, a quantificação de lignina foi realizada via espectroscopia no infravermelho próximo (NIRS). Foram selecionados os acessos que apresentaram as maiores e menores médias ao longo de todos os anos de avaliação.

Tabela 5: Lista dos vinte acessos de capim-elefante (*Cenchrus purpureus*) pertencentes ao Banco Ativo de Germoplasma de Capim-Elefante da Embrapa Gado de Leite.

Acessos de Capim-Elefante

BAGCE 5 – BRA 00058589-3

BAGCE 7 – BRA 00058591-9

BAGCE 10 – BRA 00058594-3

BAGCE 11 – BRA 00058595-0

BAGCE 12 – BRA 00058596-8

BAGCE 24 – BRA 00058607-3

BAGCE 26 – BRA 00058609-9

BAGCE 32 – BRA 00058615-6

BAGCE 58 – BRA 00058633-9

BAGCE 59 – BRA 00058634-7

BAGCE 66- BRA 00058641-2

BAGCE 67 – BRA 00058642-0

BAGCE 72 – BRA 00058647-9

BAGCE 82 – BRA 00078557-6

BAGCE 83 – BRA 00058658-6

BAGCE 91 – BRA 00058662-8

BAGCE 96 – BRA 00058667-7

BAGCE 97 – BRA 00058668-5

BAGCE 98 – BRA 00058669-3

Pioneiro – BRA 00058675-0

2.2.2 COLETA DE AMOSTRAS E EXTRAÇÃO DE RNA

As primeiras folhas jovens do primeiro entrenó foram coletadas (Figura 6) para a extração do RNA total, e todo o material usado para a coleta foi limpo com o reagente RNase Away® (Life Technologies, Carlsbad, Estados Unidos). Em seguida, as folhas selecionadas foram colocadas em tubos de 2mL com 1mL de solução RNA/later® (Life Technologies, Carlsbad, Estados Unidos), e armazenadas a -20°C.

As folhas de todos os acessos foram então maceradas em nitrogênio líquido, sendo todo o material utilizado na maceração previamente higienizado com o reagente RNase Away® (Life Technologies, Carlsbad, Estados Unidos), para eliminar qualquer contaminação da enzima RNase no processo.

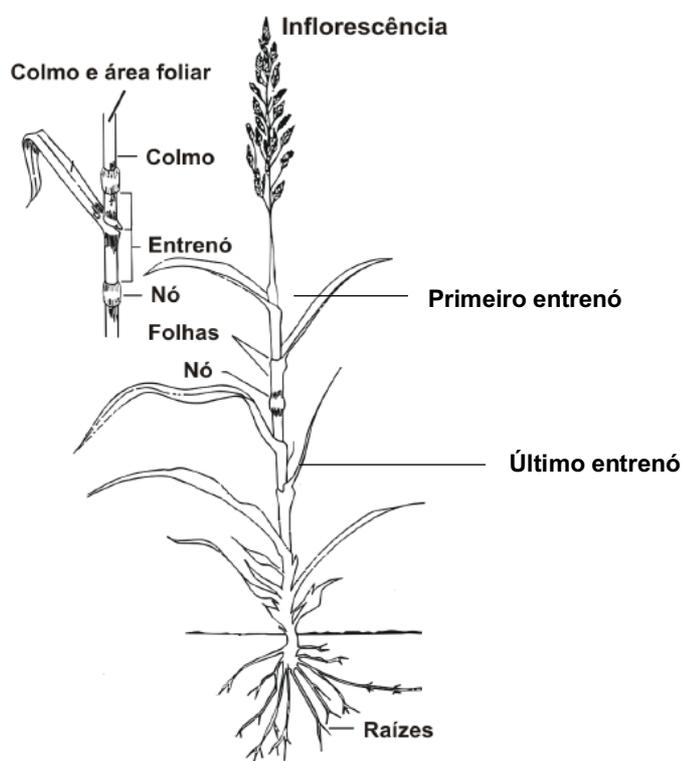


Figura 8: Morfologia típica de gramíneas. Adaptado de Ball *et al.* (2007).

O RNA das folhas jovens maceradas foi extraído utilizando o reagente TRIzol (Thermo Fisher, Waltham, Estados Unidos), seguindo o protocolo de Jordon-Thaden *et al.* (2015). Ao final do processo, o RNA extraído foi ressuspendido em 50 μ L de Nuclease-Free Water (IDT, Coralville, Estados Unidos), e aliquotado em cinco tubos de 0,5mL, com os seguintes volumes: um tubo com 3 μ L, um tubo com 2 μ L, e três tubos com 15 μ L cada.

A concentração de RNA foi estimada por leitura em Nanodrop 1000 (Thermo Fisher, Waltham, Estados Unidos), com os comprimentos de onda de 260 e 280nm.

A integridade do RNA foi estimada utilizando o equipamento Agilent 2100 Bioanalyzer e kit RNA 600 Nano (Agilent Technologies, Santa Clara, Estados Unidos), de acordo com especificações do fabricante.

Para a avaliação das porcentagens de FDA e lignina, folhas dos mesmos acessos selecionados foram coletadas, e a análise das porcentagens foi realizada seguindo a metodologia proposta por Goering e Van Soest (1967).

2.2.3 MONTAGEM DE BIBLIOTECA DE cDNA

As amostras selecionadas para o sequenciamento foram baseadas nos valores de *RIN* superiores a 6,4 e quantidade de RNA superior a 1000 ng. Para a espécie, somente as quatro amostras de maior produção de lignina e as quatro de menor produção de lignina que apresentam os melhores valores foram sequenciadas.

As etapas de conversão de RNA em cDNA e construção das bibliotecas e sequenciamento utilizando a tecnologia Illumina HiSeq2500 foram desenvolvidas pela empresa BIODOME.

As bibliotecas de DNA complementar (cDNA) foram preparadas de acordo com o protocolo Illumina TruSeq Stranded mRNA Sample Prep LT (https://support.illumina.com/content/dam/illumina/support/documents/documentation/chemistry_documentation/samplepreps_truseq/truseq-stranded-mrna_workflow/truseq-stranded-mrna-workflow-reference-1000000040498-00.pdf). Inicialmente, as amostras de RNA foram purificadas usando esferas magnéticas ligadas a oligo-dT,

que tem como função a separação do RNA mensageiro (mRNA) pela ligação a cauda poli-A contida na molécula. Em seguida, o mRNA purificado foi fragmentado e ligado a *primers* para seguir com a síntese de cDNA. A primeira fita de cDNA foi sintetizada a partir dos RNAs fragmentados, utilizando a enzima *SuperScript II Reverse Transcriptase* (Illumina). A síntese da segunda fita de cDNA foi realizada por meio do *Second Strand Marking Master Mix*, que ao final desta etapa, produziu cDNA com extremidades abruptas ligadas a esferas magnéticas. O cDNA foi então adenilado nas extremidades 3' das fitas por meio do *A-Tailing Mix* (Illumina), evitando assim que as fitas se ligassem durante a reação de ligação de adaptadores. Um nucleotídeo de timina correspondente na porção 3' dos adaptadores forneceu a ligação complementar aos fragmentos adenilados, evitando a formação de quimeras no processo. Com o cDNA ligado aos adaptadores, foi feita a hibridização dos fragmentos de cDNA na *flow cell*, realizada por meio de *Ligation Mix* (Illumina), que permite que os adaptadores encontrados nos fragmentos de cDNA se hibridizem aos adaptadores encontrados na *flow cell*. Ao final, as bibliotecas foram purificadas e amplificadas por PCR, pela ligação de *primers* aos adaptadores, na seguinte condição: 98°C por 30 s, 15 ciclos de 98°C por 10 s, 60°C por 30 s e 72°C por 30 s, seguido por 72°C por 5 min.

As bibliotecas produzidas foram avaliadas no equipamento *Bioanalyzer 2100* (Agilent Technologies) utilizando DNA 1000 chip (Agilent Technologies), para avaliação da qualidade da biblioteca. As amostras consideradas ótimas apresentaram uma banda de aproximadamente 260 pb.

2.2.4 SEQUENCIAMENTO DE mRNA (RNA-Seq)

O sequenciamento foi realizado com tecnologia de nova geração, utilizando o equipamento *HiSeq 2500* (Illumina). A metodologia baseou-se no sequenciamento por síntese, no qual cada base nitrogenada é identificada logo que é adicionada à cadeia recém-formada, por meio da enzima DNA polimerase. As amostras ligadas a adaptadores específicos são complementares a oligos presentes na superfície da *HiSeq Flow Cell v4* (Illumina) e amplificadas, produzindo *clusters* (agrupamentos) de fragmentos iguais. Em seguida, a sequência dos *clusters* foi identificada por meio do

HiSeq SBS Kit v4 (Illumina), com leituras pareadas de 100 pb (2x), de acordo com recomendações do fabricante. Essa metodologia permitiu a detecção dos nucleotídeos, complementares às sequências dos *clusters* e marcados por fluorescência, após sua incorporação a cada ciclo da reação, pela excitação da fluorescência acoplada.

2.2.5 ANÁLISE DE DADOS

2.2.5.1 Controle de Qualidade das bibliotecas

Após o sequenciamento, as bibliotecas geradas foram então avaliadas quanto à qualidade do sequenciamento. O controle de qualidade foi realizado por meio do *software* FastQC v0.11.8 (Babraham Bioinformatics, Babraham, Inglaterra), utilizando os arquivos de formato FASTQ. Para cada amostra analisada, um gráfico de Whisker foi gerado, utilizando o *Phred quality score*. O valor de *Phred* avalia a confiabilidade dos dados, atribuindo um valor de probabilidade de erro do sequenciamento para cada nucleotídeo.

2.2.5.2 Montagem do transcriptoma

O novo genoma sequenciado de *Cenchrus purpureus* (GWAHORA000000000) (YAN *et al.*, 2021b) foi utilizado como referência para a montagem do transcriptoma. O *software* Stringtie v 2.0.4 (PERTEA *et al.*, 2015) foi utilizado para a montagem do transcriptoma a partir das *reads* sequenciadas e do genoma disponível, seguindo os comandos *default* do *software*.

Para esta etapa, as *reads* foram primariamente alinhadas ao genoma por meio do *software* STAR v2.5.2b (DOBIN *et al.*, 2013). Ao final do alinhamento,

arquivos no formato BAM foram gerados e usados como *input* para a montagem do transcriptoma pelo Stringtie.

2.2.5.3 Indexação e Mapeamento de sequências

As *reads* sequenciadas foram novamente alinhadas, porém a referência foi o transcriptoma montado pelo Stringtie. Por meio do software STAR v 2.5.2b (DOBIN *et al.*, 2013), o transcriptoma foi inicialmente indexado, e em seguida as *reads* foram alinhadas a referência. Arquivos no formato BAM foram gerados com os resultados do alinhamento e usados para continuar as análises. Em todas as etapas, foram usados os comandos padrões do software.

2.2.5.4 Quantificação de transcritos

As *reads* sequenciadas foram quantificadas para calcular a expressão dos transcritos. A ferramenta Salmon v 0.7.2 (PATRO *et al.*, 2017) foi utilizada, iniciada pela indexação do transcriptoma acima montado. Em seguida, os arquivos BAM gerados pelo software STAR foram usados como *input* para calcular a quantidade dos transcritos para cada amostra. Os resultados obtidos serviram como arquivos de entrada para a análise dos genes diferencialmente expressos. Em todas as etapas, foram usados os comandos padrões da ferramenta.

2.2.5.5 Identificação de genes diferencialmente expressos

O programa R v3.6.1 (<http://www.R-project.org>), mais especificamente o pacote “DESeq2” (LOVE *et al.*, 2014) foi utilizado na identificação dos genes diferencialmente expressos em *C. purpureus*. O pacote tem como funções realizar a normalização dos dados fornecidos, garantindo melhor acurácia das amostras

analisadas, cálculo das expressões gênicas e realização de testes para avaliação das diferenças de expressão. Para corrigir possíveis erros dos testes, devido a análise de múltiplas comparações, uma etapa adicional de correção *False Discovery Rate* (FDR) é realizada.

A partir desse conjunto de dados gerados, os genes que apresentaram $p \leq 0.05$ foram considerados significativamente expressos nas amostras.

2.2.5.6 Anotação dos genes

As funções dos genes encontrados no transcriptoma de *C. purpureus* foram identificadas utilizando o software Trinotate v 3.2.0 (BRYANT *et al.*, 2017), que combina informações dos bancos de dados do KEGG (KANEHISA *et al.*, 2012) e NCBI/BLAST v 2.9.0 (ALTSCHUL *et al.*, 1990). Além de dados gênicos, o software identifica proteínas, utilizando os bancos de dados do TrEMBL/SwissProt (THE UNIPROT CONSORTIUM, 2021), Pfam (PUNTA *et al.*, 2012) e NCBI/BLAST (ALTSCHUL *et al.*, 1990), e regiões de transmembrana por meio do banco de dados do TmHMM (KROGH *et al.*, 2001). Ao final das análises, todos os dados são compilados em um único banco de dados SQLite (<https://www.sqlite.org>), e exportados para uma planilha do Excel (Microsoft Corporation, Albuquerque, Estados Unidos). Em todas as etapas de busca no banco de dados, foram seguidos os comandos padrões disponibilizados pelo software Trinotate.

2.2.5.7 Identificação dos genes diferencialmente expressos e suas anotações

As planilhas geradas pela análise dos genes diferencialmente expressos e pela anotação gênica foram combinadas, a fim de se identificar as funções dos genes considerados significativos. Essa combinação foi realizada pelo comando *merge* encontrada no programa R (<http://www.R-project.org>).

2.3 RESULTADOS E DISCUSSÃO

2.3.1 ANÁLISE BROMATOLÓGICA

Os resultados das análises bromatológicas das oito amostras selecionadas estão representadas na Tabela 6. Os valores de FDA variaram de 296 g/kg (BAGCE 5) a 374,6 g/kg (BAGCE 32), e as concentrações de lignina variaram de 25 g/kg (Pioneiro) a 439 g/kg (BAGCE 98).

Tabela 6: Análise bromatológica de oito amostras selecionadas de *Cenchrus purpureus*, com os valores de fibra em detergente ácido (FDA) e lignina, em g/kg MS, e classificação dos acessos com relação ao teor de lignina.

Acessos	FDA	Lignina	Teor Lignina
BAGCE 5	296	37,8	Baixo
BAGCE 7	328,8	43,1	Alto
BAGCE 10	328,4	38,2	Baixo
BAGCE 32	374,6	41,2	Alto
BAGCE 82	332	38,2	Baixo
BAGCE 96	331	41,4	Alto
BAGCE 98	338,3	43,9	Alto
Pioneiro	318,6	25	Baixo

Na literatura, a grande maioria dos relatos de análise bromatológica em *Cenchrus purpureus* avaliam as concentrações de FDA e lignina na planta inteira. Fernandes *et al.* (2020) estudaram seis genótipos de capim-elefante a fim de se conhecer o potencial de cada um na produção de energia. Dentre o grupo analisado, a porcentagem de FDA variou de 514,7 a 553,4 g/kg MS, enquanto a porcentagem de lignina esteve entre 117,9 e 128,4 g/kg MS. Para fins de produção de energia, porém, a análise diferenciada das concentrações de FDA e lignina é necessária. Em comparação, a produção de etanol em folhas é 10% maior do que em caules (ANDERSON *et al.*, 2008; RUEDA *et al.*, 2020), o que pode ser explicado pela

avaliação da quantidade de FDA e lignina encontrada em cada porção. Rueda *et al.* (2020) avaliaram as folhas e o caule do capim-elefante CT115 em diferentes estações, e o que foi encontrado é que folhas possuem, em média, 85 g/kg MS a menos de FDA e 24 g/kg MS a menos de lignina do que o caule. Assim, com menores valores de FDA e lignina encontrados nas folhas, a eficiência na produção de bioetanol aumenta, por consequência da diminuição da barreira física feita pela lignina no acesso de enzimas a biomassa.

Na alimentação bovina, a avaliação individual das quantidades de FDA e lignina traz informações importantes, já que folhas apresentam maior valor nutritivo do alimento do que o caule (BALL *et al.*, 2001). Habte *et al.* (2020) avaliaram 45 genótipos de capim elefante, para investigar os efeitos da interação do genótipo e ambientes no desempenho da espécie. Nas estações úmida e seca, os valores encontrados nas folhas para FDA foram entre 328,8 e 375,2 g/kg, e para lignina variaram de 40,4 e 47,4 g/kg.

2.3.2 CONTROLE DE QUALIDADE DAS BIBLIOTECAS E MONTAGEM DO TRANSCRIPTOMA

Oito bibliotecas de RNA-Seq foram sequenciadas, correspondente as quatro amostras com alta produção de lignina e quatro com baixa produção de lignina. No total foram gerados mais de 130 milhões de *reads* PE (Tabela 7), de 100 pb, com uma média de 16 milhões de *reads* por biblioteca.

Tabela 7: Total de *reads* obtidas em *Cenchrus purpureus* para cada biblioteca gerada, percentual de CG e total de *reads* sequenciadas.

Acessos	Total de reads	%CG
BAGCE 5	17.314.762	51
BAGCE 7	16.401.784	51
BAGCE 10	16.945.274	51
BAGCE 32	16.227.283	51
BAGCE 82	15.534.752	51
BAGCE 96	16.341.920	52
BAGCE 98	14.587.019	51
Pioneiro	16.960.439	51
Total	130.313.233	

A avaliação da qualidade do sequenciamento das amostras BAGCE 5 e Pioneiro pode ser vista na Figura 7. O resultado aqui obtido foi semelhante ao sequenciamento encontrado para as outras seis amostras estudadas. O eixo y representa a qualidade do sequenciamento, e quanto maior o valor encontrado, melhor é a confiança na *base call*. O valor da média da qualidade é representado pela linha central azul, e ficou na faixa de 32 a 36 para todas as amostras. Sendo assim, foi possível confirmar que o sequenciamento foi de boa qualidade, e as sequências obtidas foram usadas para as análises de dados.

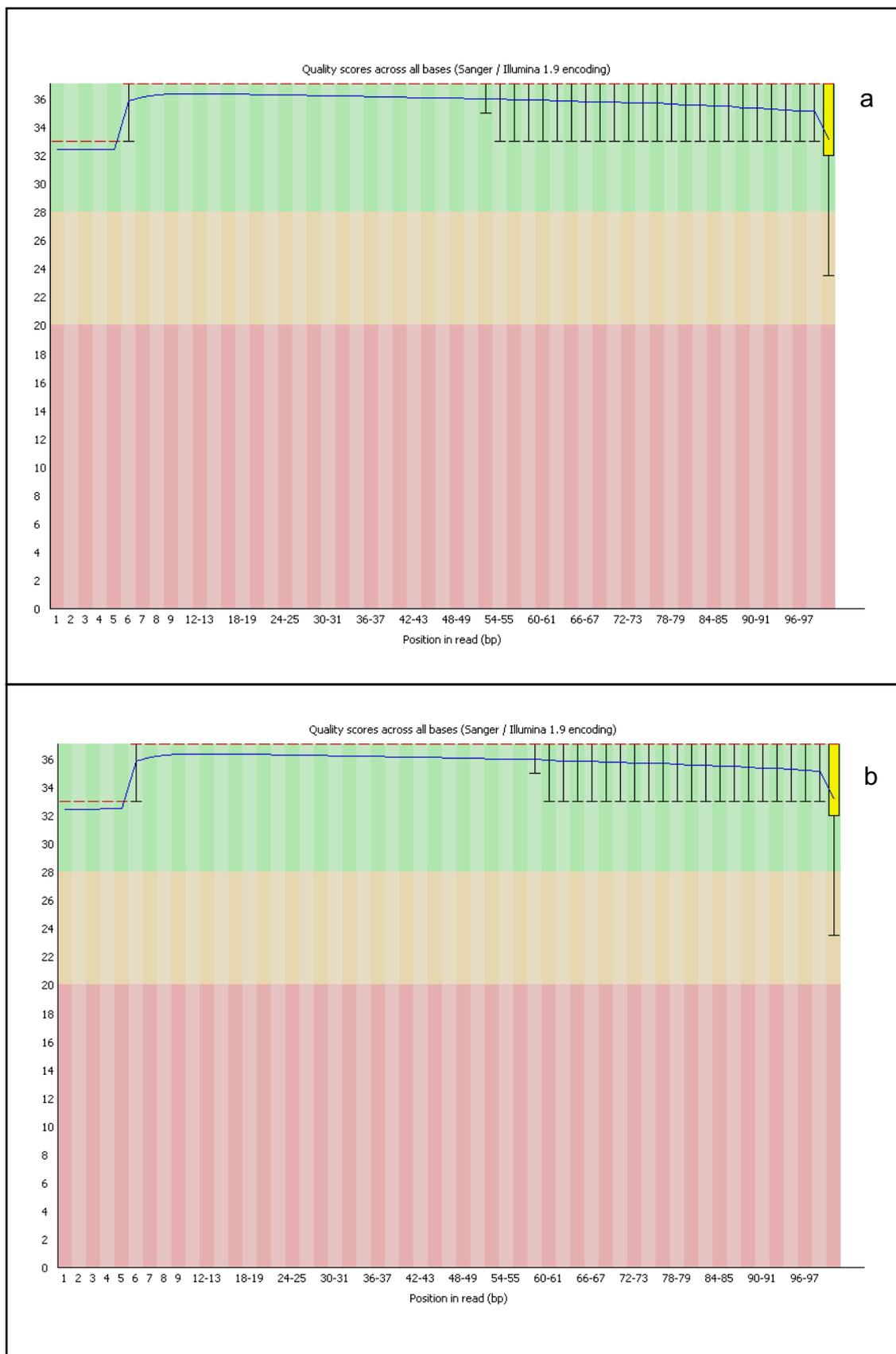


Figura 9: Gráfico de Whisker do sequenciamento do transcriptoma dos acessos de *Cenchrus purpureus* BAGCE 5 (a) e Pioneiro (b). y é o eixo dos valores de qualidades. O fundo do gráfico divide o eixo em três níveis de confiança: muito confiável (verde), qualidade razoável (laranja) e baixa qualidade (vermelho). A linha central azul representa o valor da média da qualidade. A caixa amarela representa a amplitude inter quartil (25-75%). Fonte: FastQC A Quality Control tool for High Throughput Sequence Data <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> por S. Andrews.

2.3.3 INDEXAÇÃO E MAPEAMENTO DE SEQUÊNCIAS

Para a montagem do transcriptoma, as *reads* sequenciadas foram primeiramente alinhadas ao genoma de capim-elefante (Tabela 8). O mapeamento variou de 50,07% (Pioneiro) a 52,74% (BAGCE 98) para as oito amostras analisadas.

Tabela 8: Alinhamento do sequenciamento ao genoma de capim-elefante, com o número de *reads* alinhadas e suas porcentagens.

Acessos	Reads Mapeadas (%)
BAGCE 5	8.678.424 (50,12%)
BAGCE 7	8.376.621 (51,07%)
BAGCE 10	8.852.161 (52,24%)
BAGCE 32	8.286.961 (51,07%)
BAGCE 82	7.826.406 (50,38%)
BAGCE 96	8.392.759 (51,36%)
BAGCE 98	7.693.298 (52,74%)
Pioneiro	8.491.811 (50,07%)

Em seguida, o transcriptoma foi montado pelo Stringtie utilizando os arquivos do resultado do alinhamento como *input*. O transcriptoma montado foi composto por 101.169 transcritos.

Um novo alinhamento foi realizado a fim de saber a porcentagem de *reads* mapeadas ao transcriptoma montado (Tabela 9). A porcentagem de *reads* mapeadas variou de 77,05% (BAGCE 82) a 81,7% (BAGCE 5). Os valores encontrados utilizando o transcriptoma de referência foi superior ao encontrado quando o genoma foi utilizado como referência. Este resultado é esperado, pois o transcriptoma corresponde a porção que codifica genes, chamada de éxon, enquanto o genoma é composto pelos éxons e introns, que é a parte não codificante do genoma (YI *et al.*, 2018). A partir do mapeamento realizado no transcriptoma montado, foi possível a quantificação das *reads* para cada transcrito, que foi utilizado para a análise de expressão gênica.

Tabela 9: Alinhamento das *reads* sequenciadas ao transcriptoma montado de capim-elefante, com a porcentagem correspondente a cada acesso analisado.

Acessos	% de mapeamento
BAGCE 5	81,7%
BAGCE 7	79,48%
BAGCE 10	79,61%
BAGCE 32	77,57%
BAGCE 82	77,05%
BAGCE 96	78,08%
BAGCE 98	80,43%
Pioneiro	79,26%

2.3.4 IDENTIFICAÇÃO DOS GENES DIFERENCIALMENTE EXPRESSOS E SUAS FUNÇÕES GÊNICAS

A partir da análise bromatológica realizada nas oito amostras estudadas, foram definidos dois grupos contrastes para a produção de lignina. As amostras BAGCE 7, BAGCE 32, BAGCE 96 e BAGCE 98 formaram o grupo de alta produção de lignina. As amostras BAGCE 5, BAGCE 10, BAGCE 82 e Pioneiro compõem o grupo de baixa produção de lignina. A análise dos genes DE dos contrastes alta produção x baixa produção identificaram 66.981 genes, sendo 53 (0,07%) significativos por possuírem valores de *p value* ajustado menor ou igual a 0,05. Nas amostras de alta produção de lignina, 43 genes significativos foram *up* regulados, e 10 dos genes significativos foram *down* regulados.

A anotação gênica permite a identificação de função dos genes que foram sequenciados de moléculas de DNA e RNA (ABRIL E CASTELLANO, 2019). Em capim-elefante, a anotação gênica foi realizada a partir de banco de dados de Pfam, KEGG e Gene Ontology (GO)/BLAST. 73.860 genes e suas isoformas foram anotados em pelo menos um dos três bancos de dados.

A combinação dos dados de genes DE e função gênica identificada mostrou que 46 genes com nível de significância de 0,05 (0,06%) foram anotados nos bancos de dados selecionados.

Dentre os genes anotados no banco de dados de Pfam, dois se destacam por sua relação indireta na via de biossíntese de lignina: genes da família transferase (PF02458.16) e genes da família terpeno sintase (PF01397.22 e PF03936.17). Enzimas do tipo transferase são conhecidas por transferirem grupos funcionais entre doadores e receptores, em diferentes funções celulares. Dentre essas funções, é descrito na literatura uma relação da expressão de glicosiltransferase (GT) e sacarificação no processo de biorefinaria. A enzima GT é conhecida pela biossíntese de xilano, um polissacarídeo presente na fração de hemicelulose da parede celular secundária (BROWN *et al.*, 2007; WHITEHEAD *et al.*, 2018). Este polissacarídeo, por sua vez, é um dos responsáveis pela recalcitrância de plantas (WHITEHEAD *et al.*, 2018) e, por consequência, baixa eficiência no processo de sacarificação (BHATIA *et al.*, 2016; KO *et al.*, 2020). Estudos envolvendo a enzima GT, xilano e sacarificação revelaram que a supressão de GT em plantas diminui a concentração de xilano, aumentando assim a eficiência de sacarificação (LEE *et al.*, 2009; BISWAL *et al.*, 2015; KO *et al.*, 2020). A família terpeno sintase é responsável pela produção de uma diversidade de terpenos/terpenoides, metabólito secundário envolvido em processos de estresses biótico e abiótico, e de interação planta-patógeno (BONCAN *et al.*, 2020; NAGEGOWDA E GUPTA, 2020; JAHANGEER *et al.*, 2021). Um estudo publicado por Jozwiak *et al.* (2020) revelou que a produção de terpenos, assim como a sua modificação, pode ter interferência de enzimas presentes na via de biossíntese das paredes celulares primárias e secundárias. Enzimas relacionadas a celulose sintase são, como descritas no trabalho, “sequestradas” da construção da parede celular para a produção de diversos conjuntos de terpenos relacionados a defesa da planta.

Por meio do banco de dados KEGG, alguns genes apareceram mais de uma vez a partir das sequências analisadas. A proteína FAR-1 (*Fatty Acil-CoA Reductase 1*) (*ath*:AT3G06250; *ath*: AT5G18960) está relacionada ao mecanismo de defesa de plantas quando invadidas por nematoides (DING *et al.*, 2019; LI *et al.*, 2020). A enzima fosfoserina fosfatase (*ath*:AT1G18640) é uma etapa importante de fosforilação da serina, intermediário importantes em vias do metabolismo de plantas (WULFERT E KRUEGER, 2018). Além desta função, a enzima tem um possível papel

na resposta a estresse salínico em *Brassica juncea* (PURTY *et al.*, 2017) e no crescimento de plantas dependente da luz e de açúcar (WULFERT E KRUEGER, 2018).

Assim como visto no banco Pfam, foi possível identificar pelo GO/BLAST genes envolvidos na biossíntese de terpenos/terpenoides (GO:0016114). Pode-se destacar dois outros genes identificados neste banco de dados: genes responsáveis pela biossíntese de esterol (GO:0016126) e de *L*-serina (GO:0006564). Esteróis são compostos que possuem funções diversas e essenciais em eucariotos, principalmente na composição e manutenção da permeabilidade e fluidez de membranas celulares (HARTMANN, 1998; CLOUSE, 2002, ROGOWSKA E SKAZIEL, 2020). Esta última função torna-se importante no processo de polimerização de lignina, pois os monômeros são produzidos no citoplasma das células, e precisa ser transportado para fora da célula, através da membrana plasmática (BOIJA E JOHANSSON, 2006). Já a *L*-serina é um aminoácido predominantemente sintetizado a partir das vias de glicosilação e fosforilação, a partir do aminoácido serina. Além das funções associadas a serina descrita acima, sua concentração em células foi relacionada na produção de biomassa e recalcitrância. Zhang *et a.* (2019) avaliaram a superexpressão de serina hidroximetiltransferase (SHMT), enzima que catalisa a conversão de *L*-serina em glicina. No estudo, foi possível identificar que plantas transgênicas para a super expressão do gene SHMT em *Populus* (*PtSHMT2*) tiveram um aumento da produção de biomassa e açúcares e diminuição da concentração de lignina, por desencadear a expressão de MYB4, conhecido por ser regulador na produção de lignina.

Com este resultado, não foi possível identificar genes já descritos e conhecidos envolvidos na via de lignina em *Cenchrus purpureus*. Alguns fatores podem ter contribuído para a ausência destes genes nos dados finais, como a porção da planta coletada e o agrupamento das amostras realizadas. Na descoberta de expressão de genes presentes na via de lignina, a porção vegetal escolhida geralmente é o caule, pois a lignina auxilia na sustentação da planta, principalmente em células vasculares localizadas no caule (YU *et al.*, 2020). Um exemplo é o trabalho de Zhang *et al.* (2020), que compararam a expressão de genes de lignina em caules de capim-elefante, com maturação variando de 40 a 120 dias. Porém, no delineamento experimental deste trabalho, preconizou-se a escolha de folhas jovens

dos 20 acessos promissores, pois essa porção vegetal é a mais utilizada como forragem e biomassa (LYONS *et al.* 1996; BALL *et al.*, 2001). Folhas jovens não sofrem influência do corte, e assim como relatado por Rocha *et al.* (2019), a interação de genótipo e corte influenciam significativamente na produção de FDA, digestibilidade da biomassa e teor de lignina nos genótipos de capim-elefante. Em contrapartida, folhas novas são conhecidas por produzir 10% mais etanol do que o caule (ANDERSON *et al.*, 2008). O agrupamento dos genótipos com relação a produção de lignina foi baseado no grupo bem estudado por 6 anos pelo Programa de Melhoramento Genético de Capim-Elefante da Embrapa Gado de Leite, combinado a análise bromatológica realizada no momento da coleta e a variação genética com relação a FDA e quantidade de lignina entre os genótipos de capim-elefante apresentada por Rocha *et al.* (2019).

A avaliação bromatológica de capim-elefante ao longo dos anos evidenciou que o teor de lignina e FDA se mantiveram constantes em genótipos específicos. A exemplo, BAGCE 5 e Pioneiro tiveram um teor de lignina baixo constante, enquanto BAGCE 96 e BAGCE 98 foram caracterizados com alto teor de lignina na planta. Por meio desta informação, um novo agrupamento foi realizado a fim de se identificar genes expressos que podem elucidar a diferença encontrada entre os genótipos com relação a concentração de lignina. Os resultados da análise de genes DE combinado a anotação gênica deste novo grupo identificaram 1.950 genes DE, com funções gênicas anotadas nos bancos do Pfam, KEGG e GO/BLAST, a nível de significância de 0,05%. Destes, 1.214 genes foram *up* regulados, e 736 genes foram *down* regulados. Por meio dos bancos de dados, foi possível identificar genes da família MYB (PF13837.7), fator de transcrição conhecido por regular a biossíntese de lignina em plantas (ZHONG *et al.*, 2009; WANG E DIXON, 2012; NAKAMO *et al.*, 2015; YU *et al.*, 2020), e peroxidase (PF00141.24), enzima presente na polimerização da lignina (BOERJAN *et al.*, 2003). Outros genes que chamam a atenção são os da celulose sintase (PF03552.15), família glicosil hidrolase (PF01373.18) e expansina (*osa*:4336479). A expressão destes genes pode ter relação com a escolha de coletar folhas jovens das plantas. Neste estágio inicial de desenvolvimento, as células sofrem alongamento, e as paredes celulares primária e secundária estão em formação e remodelamento (SAMPEDRO E COSGROVE, 2005; MORGAN *et al.*, 2016; PURUSHOTHAN *et al.*, 2020).

Com este novo agrupamento, foi possível anotar funções gênicas ainda não identificadas no primeiro agrupamento. Porém, ainda não foi possível encontrar os genes da via de biossíntese de lignina nos genótipos de capim-elefante avaliados. Por isso, foi realizada uma prospecção dos genes já conhecidos de lignina no transcriptoma de *C. purpureus*, focando principalmente no agrupamento BAGCE 5/Pioneiro (baixo teor de lignina) X BAGCE 96/BAGCE 98 (alto teor de lignina). A Figura 8 mostra o nível de expressão das famílias gênicas chaves na biossíntese de lignina nos genótipos avaliados, representado em *heatmap*. A família gênica de maior expressão é a família COMT, com maior expressão no genótipo BAGCE 98 (aproximadamente 300 TPM). Na via de biossíntese de lignina, esta enzima está envolvida na metilação do precursor da unidade S (BONAWITZ E CHAPPLE, 2010; CESARINO *et al.*, 2016). Dada a alta expressão, foi testada a hipótese de que este resultado esteja influenciando nos resultados de expressão de outras famílias gênicas. Ao retirar a família COMT da análise (Figura 9), pode-se observar que outras famílias gênicas surgiram com expressão significativa, como a família PTAL no genótipo BAGCE 98, e o gene CCoAOMT1b1, que pertence a família gênica CCoAOMT, nos genótipos BAGCE 96 e BAGCE 98. A enzima PTAL está presente na via alternativa da biossíntese de lignina, que se inicia pelo aminoácido tirosina (ROSLER *et al.*, 1997; WATTS *et al.*, 2006; BARROS *et al.*, 2016; CESARINO *et al.*, 2016). Já as enzimas da família CCoAOMT têm função de metilação no precursor do coniferaldeído, importante molécula para a formação das unidades G e S (BONAWITZ E CHAPPLE, 2010; CAROCHA *et al.*, 2015; CESARINO *et al.*, 2016).

A fim de se utilizar os conhecimentos de expressão gênica de *Cenchrus purpureus*, focando na característica de produção de lignina, no melhoramento da espécie, novas investigações são necessárias, como a avaliação da expressão gênica em outras porções vegetais, e a escolha de acessos que apresentem uma produção de lignina mais discrepante do que foi encontrada neste trabalho. As análises realizadas no sequenciamento do RNA de folhas jovens da espécie pouco revelaram genes chave ligados a característica de produção de lignina, mesmo restringindo a genótipos que, ao longo dos anos de avaliação, mantiveram constante a produção.

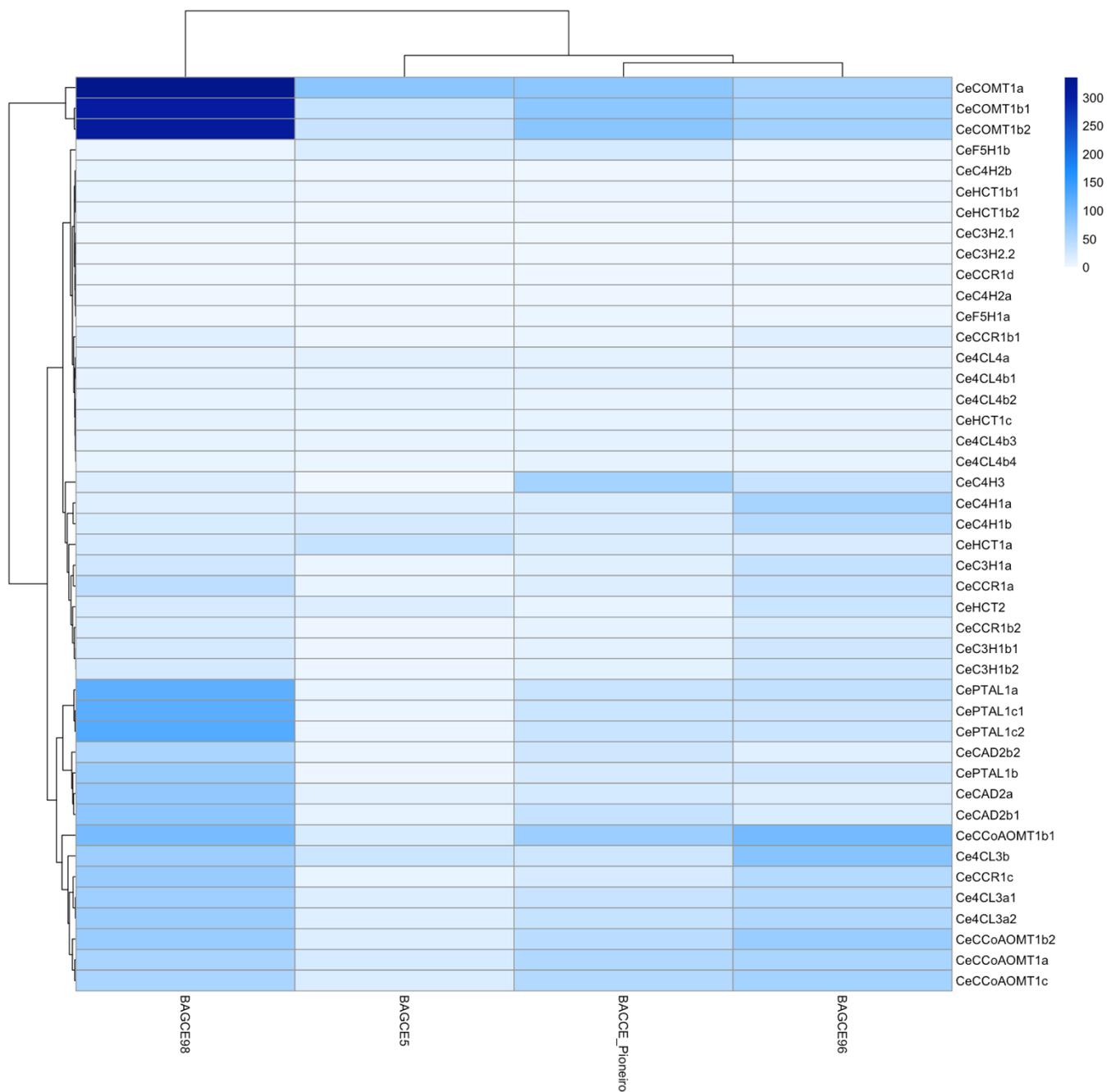


Figura 10: Nível de expressão gênica de genes importantes na via de biossíntese de lignina nos acessos BAGCE 96, BAGCE 98, BAGCE 5 e Pioneiro. No canto superior direito, pode-se observar o indicador de nível de expressão, em TPM (*transcripts per million*). Quanto mais escuro for o azul, maior é a expressão deste gene.

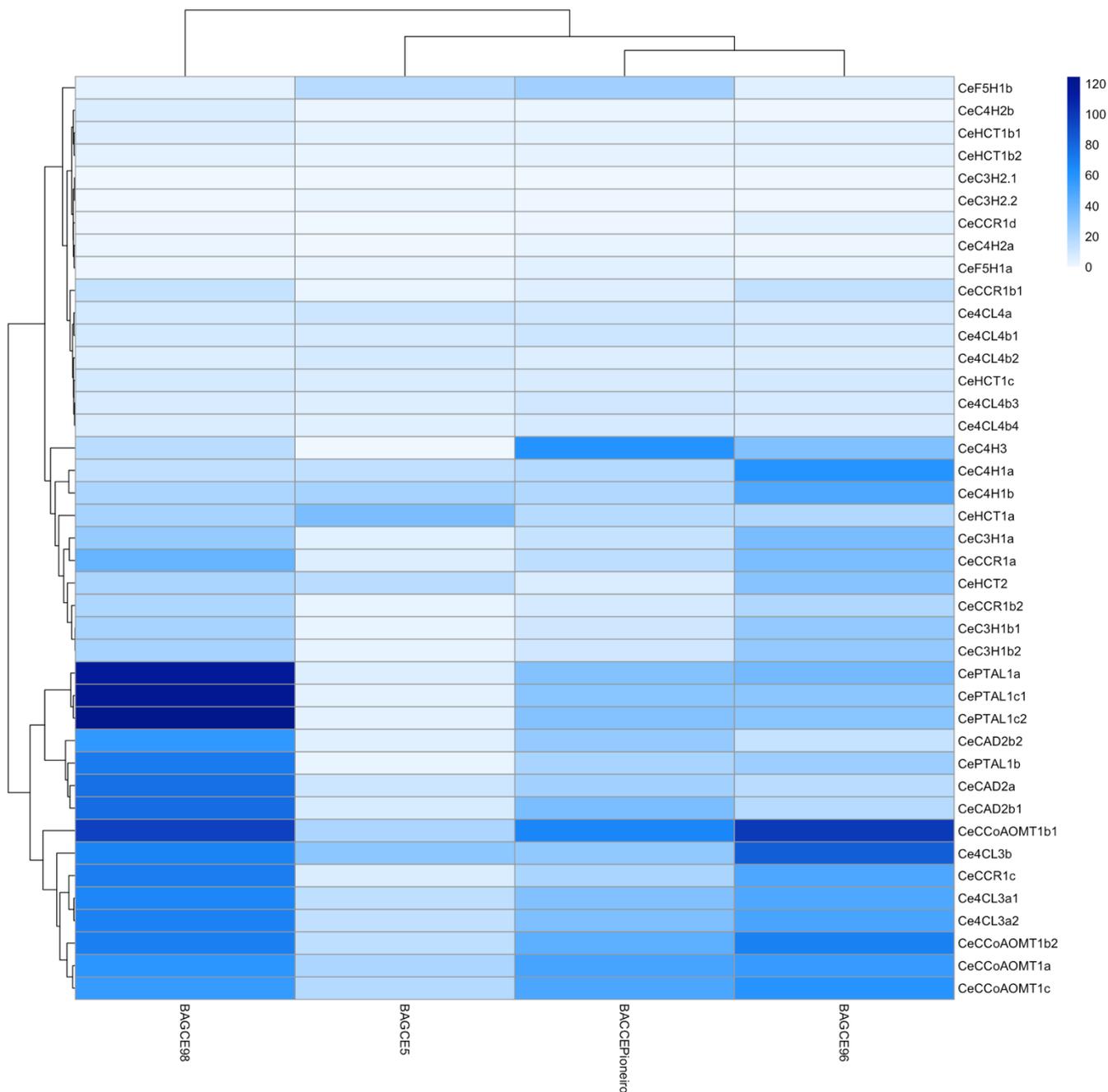


Figura 11: Nível de expressão gênica de genes importantes na via de biossíntese de lignina, excluindo-se os genes da família COMT, nos acessos BAGCE 96, BAGCE 98, BAGCE 5 e Pioneiro. No canto superior direito, pode-se observar o indicador de nível de expressão, em TPM (*transcripts per million*). Quanto mais escuro for o azul, maior é a expressão deste gene.

2.4 CONCLUSÃO

O transcriptoma montado baseando-se no sequenciamento de RNA de oito acessos de *Cenchrus purpureus* não se mostrou informativo acerca de genes envolvidos na via de biossíntese de lignina. Genes diferencialmente expressos de quatro acessos bem caracterizados acerca da produção de lignina revelou genes importantes na construção da parede celular. A prospecção de genes conhecidos na via de lignina nas quatro amostras selecionadas evidenciou uma maior expressão do gene COMT, e em seguida, genes da família CCoAOMT e PTAL. O conhecimento de genes desta via ainda precisa ser mais bem estudado para poder utilizar como ferramenta no melhoramento genético da espécie, para fins de alimentação e produção de energia.

CAPÍTULO 3: IDENTIFICAÇÃO DE MARCADORES MOLECULARES EM *U. humidicola* E *C. purpureus*

3.1 INTRODUÇÃO

Em programas de melhoramento, a seleção dos melhores materiais, que poderão se tornar novas cultivares, podem ser feitas por meio de diferentes metodologias. Inicialmente, a seleção de materiais era baseada principalmente em características morfológicas, como estrutura da semente, cor da flor e hábito de crescimento. Com a descoberta de ferramentas moleculares, foi possível associar algumas características de interesse a marcadores moleculares, o que contribuiu para uma seleção mais precisa de materiais superiores, sem a influência do ambiente e do estágio de desenvolvimento vegetal (NADEEM *et al.*, 2018). A seleção por marcadores moleculares foi a principal responsável pela seleção de cultivares de espécies como arroz (MACKILL *et al.*, 1999; GARRIDO-CARDENAS *et al.*, 2018), milho (ORTIZ, 2010; SUWARNO *et al.*, 2015) e trigo (LANDJEVA *et al.*, 2007; GARRIDO-CARDENAS *et al.*, 2018), a partir da investigação de polimorfismos encontrados a nível molecular (NADEEM *et al.*, 2018).

Dentre os marcadores moleculares, *Simple Sequence Repeat* (SSR), mais conhecido como microssatélites, são marcadores co-dominantes e com alto nível de polimorfismo (ADAMS *et al.*, 2006; RODRIGUEZ *et al.*, 2010; WANG *et al.*, 2012; HSU *et al.*, 2013; ZHANG *et al.*, 2014; LIU *et al.*, 2020a). A reunião dessas características o tornou a principal ferramenta em estudos de investigação de diversidade genética e heterose (WU *et al.*, 2013; NADEEM *et al.*, 2018), fluxo gênico (LIU *et al.*, 2020a; LU *et al.*, 2021), mapeamento de QTL (*Quantitative Trait Loci*) (NADEEM *et al.*, 2018; SU *et al.*, 2018; CHOI *et al.*, 2019) e no *fingerprint* do DNA, importante ferramenta na identificação de cultivares.

O *fingerprint* do DNA permite a identificação precisa de uma cultivar através da utilização de marcadores moleculares, assegurando a proteção intelectual e garantindo a pureza do material que está em comercialização (SATTURU *et al.*, 2018; NIE *et al.*, 2019). A ferramenta se baseia na identificação de características únicas presentes em cultivares por meio de marcadores moleculares, como forma de caracterizar as variedades de forma mais precisa e confiável, se comparado as observações de características morfológicas (IQBAL *et al.*, 2019; IQBAL *et al.*, 2021; TIAN *et al.*, 2021).

Urochloa humidicola e *Cenchrus purpureus* são espécies forrageiras comercializadas no Brasil como fonte de alimentação de bovinos, e que ao longo dos anos, cultivares foram desenvolvidas para atender as demandas do mercado. Apesar da importância, existem poucos trabalhos disponíveis relacionados a estudos moleculares em ambas as espécies. O primeiro painel de marcadores microssatélites para *U. humidicola* foi descrito por Jungmann *et al.* (2009) a partir de transferência de marcadores de outras espécies de *Urochloa*. A descoberta de um número maior de marcadores microssatélites pode trazer contribuições em etapas futuras no programa de melhoramento da espécie, na identificação de características importantes no desenvolvimento de cultivares. Em capim-elefante, Azevedo *et al.* (2012) fizeram a transferibilidade de microssatélites desenvolvidos em milho para a espécie, que foram importantes em estudos do programa de melhoramento desenvolvido pela Embrapa Gado de Leite. A Embrapa foi responsável pelo lançamento de cultivares importantes como Pioneiro, BRS Kurumi, BRS Canará e BRS Capiáçu. Essas cultivares se destacam pela alta produtividade e qualidade da forragem, além de ser uma alternativa como silagem (PEREIRA *et al.*, 2016a). Porém, ainda não está disponível um painel de *fingerprint* do DNA para essas cultivares, que é muito importante para garantir a pureza dessas cultivares no mercado forrageiro.

O objetivo deste estudo é descobrir novos marcadores SSRs a partir do sequenciamento do RNA de *Urochloa humidicola* e *Cenchrus purpureus*, e identificar um painel *fingerprint* de DNA para caracterização precisa de quatro cultivares de capim-elefante (BRS Kurumi, BRS Canará, BRS Capiáçu e Pioneiro).

3.2 MATERIAL E MÉTODOS

3.2.1 IDENTIFICAÇÃO DE MICROSSATÉLITES E DESENHOS DE *PRIMERS*

A partir dos transcriptomas montados das espécies *U. humidicola* e *C. purpureus*, regiões de microssatélites foram identificadas por meio da ferramenta MISA v 1.0 (BEIER *et al.*, 2017). Essa ferramenta é usada por fornecer a localização dos microssatélites, assim como a composição das repetições e seus intervalos.

Após a identificação dos SSRs, *primers* foram desenhados a partir do programa Primer3 v 2.3.4 (UNTERGASSER *et al.*, 2012), utilizando critérios como tamanhos de *primers* variando entre 18 e 25 pb, e tamanho dos produtos variando de 100 a 400 pb.

3.2.2 TESTE DE AMPLIFICAÇÃO DOS *PRIMERS* DESENHADOS

Cinquenta *primers* de microssatélites foram escolhidos para o teste de amplificação para cada uma das duas espécies. Os *primers* selecionados foram aqueles que apresentaram sequência de repetições de di-, tri- e tetranucleotídeos, com o produto de amplificação entre 100 e 400pb.

Quatro amostras de *Urochloa humidicola* (22A, 23A, 25A e 28A) e quatro amostras de capim-elefante (BAGCE 11, BAGCE50, BAGCE 57 e BAGCE 69) foram testadas a fim de se identificar quais pares de *primers* funcionariam para cada uma das espécies.

A reação em cadeia de polimerase foi realizada em um volume final de 20 μ L como descrito: 1X GoTaq reaction buffer, 0,5 μ M de cada *primer*, 3mM MgCl₂, 0,4mM dNTP (Promega, Wisconsin, Estados Unidos), 1 unidade de GoTaq Flexi DNA Polimerase (Promega, Wisconsin, Estados Unidos) e 45 ng de DNA genômico. As

reações em cadeia da polimerase foram realizadas em termociclador Veriti (Life Technologies, Califórnia, Estados Unidos) com o seguinte perfil: desnaturação inicial a 95°C (15 minutos), seguido de 5 ciclos a 94°C (30 segundos), temperatura de ligação de 57°C (90 segundos) e 72°C (1 minuto), com decréscimo de 1°C na temperatura de ligação por ciclo; 25 ciclos a 94°C (30 segundos), temperatura de ligação de 52°C (90 segundos) e 72°C (1 minuto); e um ciclo final de extensão de 60 minutos a 60°C.

Em seguida, os produtos de amplificação foram submetidos à eletroforese em gel de agarose a 2% por 2 horas e 30 minutos a uma voltagem de 120 V. Após o tempo de corrida, o gel foi corado em brometo de etídio por 30 minutos, e as bandas foram reveladas por meio de luz UV acoplada ao aparelho EagleEye.

3.2.3 PAINEL ESPECÍFICO PARA IDENTIFICAÇÃO DE CULTIVARES DE *Cenchrus purpureus*

Dos 50 pares de *primers* testados para capim-elefante, 21 apresentaram bandas polimórficas entre as quatro amostras inicialmente testadas. A partir desta informação, os *primers* foram testados para identificação de polimorfismo em 21 acessos do BAGCE e cultivares desenvolvidas pelo Programa de Melhoramento de Capim-Elefante da Embrapa (BAGCE 58, BAGCE 30, BAGCE 2, BAGCE 105, BAGCE 60, BAGCE 57, BAGCE 1, BAGCE 18, BAGCE 56, BAGCE 71, BAGCE 53, BAGCE 67, BAGCE 103, BAGCE 7, BAGCE 64, BAGCE 68, BAGCE 70, BRS Canará, BRS Kurumi, Pioneiro e BRS Capiaçú). Os acessos do BAGCE foram escolhidos de acordo com o grau de diversidade previamente avaliado (AZEVEDO *et al.*, 2012) de forma a contemplar o máximo de diversidade possível.

A reação em cadeia de polimerase foi realizada em um volume final de 20 µL como descrito: 1X GoTaq reaction buffer, 0,5 µM de cada *primer*, 3mM MgCl₂, 0,4mM dNTP (Promega, Wisconsin, Estados Unidos), 1 unidade de GoTaq Flexi DNA Polimerase (Promega, Wisconsin, Estados Unidos) e 45 ng de DNA genômico. As reações em cadeia da polimerase foram realizadas em termociclador Veriti (Life Technologies, Califórnia, Estados Unidos) com o seguinte perfil: desnaturação inicial a 95°C (15 minutos), seguido de 5 ciclos a 94°C (30 segundos), temperatura de ligação

de 57°C (90 segundos) e 72°C (1 minuto), com decréscimo de 1°C na temperatura de ligação por ciclo; 25 ciclos a 94°C (30 segundos), temperatura de ligação de 52°C (90 segundos) e 72°C (1 minuto); e um ciclo final de extensão de 60 minutos a 60°C.

Os produtos de amplificação foram submetidos a eletroforese em gel de poliacrilamida a 12% por cinco horas a uma voltagem de 500 V. Posteriormente o gel foi corado com nitrato de prata.

3.3 RESULTADOS E DISCUSSÃO

3.3.1 TESTE DE AMPLIFICAÇÃO

Cinquenta pares de *primers* foram selecionados dentro de cada uma das duas espécies estudadas para realizar o teste em quatro amostras de *Urochloa humidicola* e quatro amostras de *Cenchrus purpureus*.

Em *Urochloa humidicola*, 42 pares de *primers* (84%) (Figura 12) amplificaram nas quatro amostras testadas (Tabela 10). A partir desse resultado, aqueles que apresentaram polimorfismo nas amostras foram considerados promissores na identificação de cultivares da espécie.

Tabela 10: Lista dos 42 marcadores microssatélites que amplificaram nas quatro amostras de *Urochloa humidicola*, com suas sequências *forward* e *reverse*, o motivo de repetição e o tamanho de amplificação.

Nome Marcador	Primer forward	Primer reverse	Motivo repetição	Tamanho (pb)
RNA-Bra 01	GCCTTCTCTTCGACGACCTC	CCGACGTTGTTTCATCTCCCA	(CGC) ₆	190
RNA-Bra 02	GCAGCAGTTCGAGACCATGA	AGTTTCTCCGTGTAGCTCGC	(GGC) ₆	141
RNA-Bra 03	CAAGAAGAGGGTGGAGTGGC	ACTCCGAGTCAATTCGAGCG	(GGC) ₆	201
RNA-Bra 04	ATCTCGCACCTGAGCTCTA	GATATGGCGTCGCGGATCAT	(CAC) ₅	230
RNA-Bra 05	AACTCCTTTTGCCACACCCA	GTAGACGAGGATCTGCTCGC	(GCG) ₅	264
RNA-Bra 06	GGGCCAAAGGAAAGTTGTGG	GGTCTAGGAAGATCGGCAGC	(TTC) ₅	182
RNA-Bra 07	GACCTACAACATCACCGGCT	CCTTCTTCGTCCCCACCATC	(GGC) ₅	155
RNA-Bra 08	TCGGCTACGACATCAAGGTC	GACTACGCTACCGACGGAAG	(CGG) ₅	250

RNA-Bra 09	TTGGGGAGGATGCCGATAGA	CTCTCCAGCTGGGGAAAGTG	(CGC) ₅	221
RNA-Bra 10	ACGAGCAGATCCTTGACGAC	TCCTGGATGAACGCCTTCAC	(GCC) ₅	177
RNA-Bra 11	CCGCTCTACCCGAGATCCTA	GTAGAACTCCTCCACCTGCG	(GCG) ₅	329
RNA-Bra 12	GCGCATCGAGATCTACACCA	GCCACCGTTCAAGAAGGAGA	(GGC) ₆	362
RNA-Bra 13	GAATCTTCCACGCGTGATGC	CGTCTCCTTCTCCTCCTCCA	(AAG) ₇	270
RNA-Bra 14	GCGTTGCTGAACAGGAAGTG	TGCCATGCATTAGCTGTGGA	(GAA) ₅	177
RNA-Bra 15	CTAACCCCTAATCCCGTCGCC	GCCCATAGCGAGCAAGATCT	(GGC) ₅	308
RNA-Bra 16	AGAATAGATGACGCCGCTCG	AGAAGAAGGAGGTGCCGTTG	(TCC) ₆	272
RNA-Bra 17	CGAGCAGACCTACGGCAAG	AGCAGCTAGGAAGACGAGGA	(CGA) ₅	177
RNA-Bra 18	GTGCGTGCAAGCGATAAGAG	GATCACGCTCGATTTGGTGC	(CCG) ₅	207
RNA-Bra 19	CGTACCAGATGGAGCAGTCC	TGGAAGGCAGGGATCCTACA	(GCG) ₅	280
RNA-Bra 20	CTCCAGCTGCTATGCTCCTC	GACCCCGGTATGACCATGTC	(CGC) ₅	285
RNA-Bra 23	CTCCCCTTCCACCTCGAAAC	GTGACACCTTGGACCAGGAG	(CGC) ₇	276
RNA-Bra 24	GCTCCTCCAAGAAGAGCCAG	ACTCACAGTTCCACCTCCT	(CAG) ₅	168
RNA-Bra 27	GGCACGAGAGAAGGTTCCAA	CCCCATGTTGTACAGCAGGT	(CAC) ₅	144
RNA-Bra 28	CAAGAGGAGGAAGGGTCGTC	TGTCGTACATGTGGACGAGC	(CGG) ₅	238
RNA-Bra 29	CAGTTCAGCAGCCTCAGAA	ATTCTGTGGCGTTTGCACAC	(AGC) ₅	268
RNA-Bra 30	GGATTAGGTGGGAGGGAGGT	CTCTCCCCTCTCACTCAGCT	(AGC) ₅	237
RNA-Bra 31	GGGTGGTGGAGTAAGGAAGC	CGGATGAAGTGCTTGTACGC	(GCG) ₆	310

RNA-Bra 32	ATTCTTCTTCGCCCCGTCAGG	GGCCGTCCATTTTATTCCGC	(GCC) ₅	317
RNA-Bra 33	CGACTCCGTCTGGAATTGGT	ATGTTCTTGGAGAGGCAGGC	(GCG) ₅	125
RNA-Bra 34	CTCTCCCCTCTCACTCAGCT	GGATTAGGTGGGAGGGAGGT	(GCT) ₅	237
RNA-Bra 35	CTCTTCGCGTCGTCCTACTC	CAGGTTCTTGAGCTCCTGCA	(GGC) ₅	210
RNA-Bra 36	AACTGGGCTCAATCGGACAG	ATCTGGGTAGCAGGACTGGT	(CAT) ₅	292
RNA-Bra 37	TTCCCCTGCTCGTGATCAAC	TAGACGCTGGCCTGCTACTA	(CCA) ₅	325
RNA-Bra 38	CTCCTCTTCCTCACAGCAGC	CTGGTAAGTCTGGCTTGGCA	(CAG) ₈	333
RNA-Bra 39	ATCTCACAGTCCGACCTCCA	CGCTCATGTTTATGTTGGGC	(CAG) ₉	143
RNA-Bra 41	CGTCCTACTCGCACTCCATC	TTACTGCCAAGGGGACGAAC	(CAG) ₅	366
RNA-Bra 42	AGGAGTTGAGCTCAGCACAC	GCAATGCTCCTCCTGAAGGT	(GCC) ₅	234
RNA-Bra 44	AAGGCGCAAATCAGCTTGG	CAGCGTGATCTCGACGATGA	(CCT) ₇	363
RNA-Bra 45	AAGAAGAAGGGCGACAAGGG	TCGACGTCCTCCTCATCCTT	(GAC) ₆	254
RNA-Bra 46	CCGAGCTCAAGGTGAAGGAG	TAGTGGTACAACAGCCTGCG	(GC) ₆	354
RNA-Bra 48	TTCCTCCACGACAACGACAG	AAAAGGACGGCTCCCATCAG	(AGC) ₅	321
RNA-Bra 49	GAGGAAGGATGGGTGGTGTG	TGATGGAAGTTGAGCTGCGT	(GCG) ₅	365

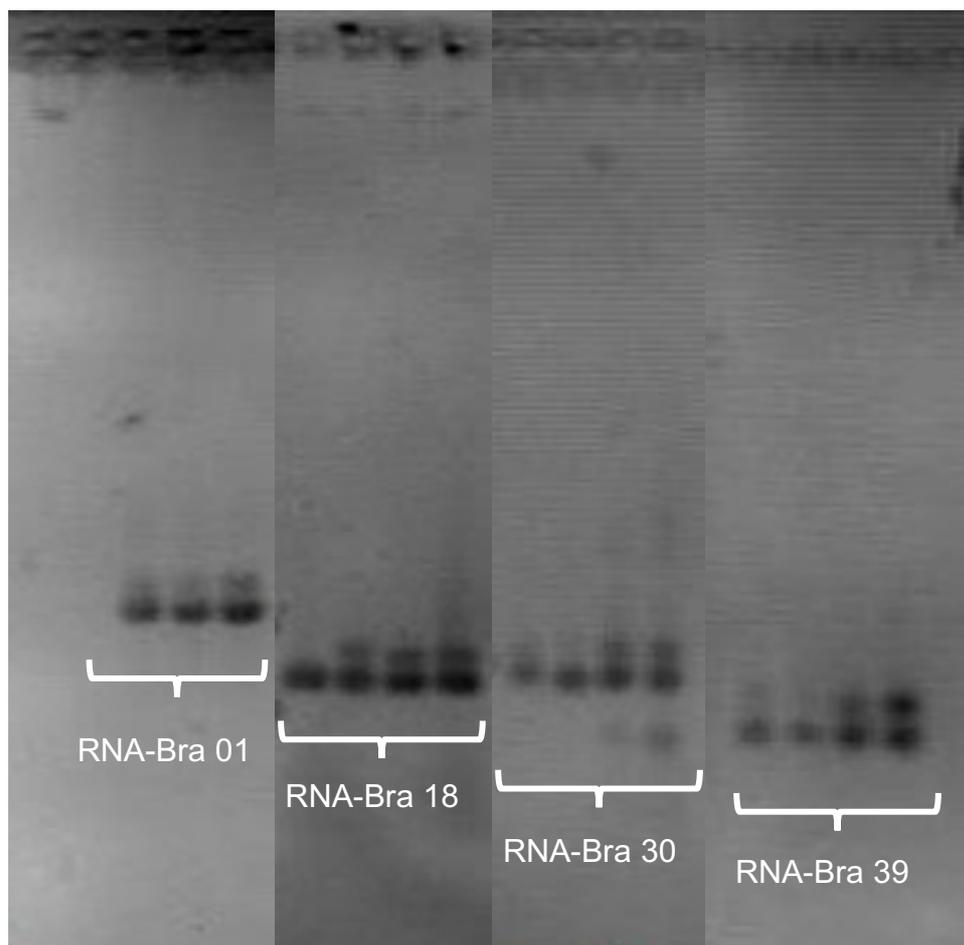


Figura 12: Géis de agarose com os *primers* RNA-Bra 01, RNA-Bra 18, RNA-Bra 30 e RNA-Bra 39 nas amostras 22A, 23A, 25A e 28A de *U. humidicola* inicialmente testadas.

Quarenta e sete pares de *primers* (94%) foram amplificados nas quatro amostras de *Cenchrus purpureus* inicialmente testadas (Figura 13). Destes, 21 *primers* apresentaram polimorfismo entre as amostras selecionadas, e assim foram considerados promissores para a identificação de cultivares na espécie (Tabela 11).

Tabela 11: Lista dos 47 marcadores microssatélites que amplificaram nas quatro amostras de *Cenchrus purpureus*, com suas sequências *forward* e *reverse*, o motivo de repetição, o tamanho de amplificação, e os *primers* selecionados para a identificação de *fingerprint* em cultivares.

Nome marcador	<i>Primer forward</i>	<i>Primer reverse</i>	Motivo repetição	Tamanho (pb)	Identificação <i>fingerprint</i> em Cultivares
RNA-CE 01	TTGCGATGCACCACAACCTTG	GGCAGCAGGTGAATCTTCCT	(GAT) ₅	188	Sim
RNA-CE 02	TACACCACCTCCTTAGCCGA	TGGTTGATAGCCGTCCATCG	(TGC) ₅	360	Sim
RNA-CE 04	GGCCTCTCCTTTTCGCTCTTT	TTTGGCCGGTTGCTAGGATT	(TG) ₇	267	Sim
RNA-CE 05	GGTCTAATGCCGGATCAGGG	GCAATGCCCATGCTAGATGC	(GCC) ₅	125	Sim
RNA-CE 06	CGGCGACATGAAGTCCTTCT	GAAGGGATGAACGCGATTGC	(GCG) ₇	399	Não
RNA-CE 07	TTCTCACATCAGCTCGCTGG	CATTGGAGAGACGGAGCGAG	(CCG) ₇	221	Não
RNA-CE 08	TCCTCCCGCTTTACCCAAAC	TTCTCGGCATCTGCAACT	(CGG) ₇	304	Não
RNA-CE 09	GTCTACAACACCTTCGGCGA	GTCGACCATCCGTTGTACT	(GAC) ₅	324	Não
RNA-CE 10	TCCTCCTCTCCCTCTCAAGC	CACATCACCAGCCAAGGAGT	(CTC) ₆	196	Não
RNA-CE 11	AGCAGGGGAGGAGAGGAAAT	GAGCACCACGAACAGGATCA	(CAG) ₇	119	Não
RNA-CE 12	GAGTGGATGTTGAGGCAGCT	AAACAGGCACGCTCTAGCTT	(GAT) ₅	246	Sim
RNA-CE 13	GTATGCACGCCAATTGCCAT	ACCACACAACAGCCGAGAAA	(TG) ₇	367	Não
RNA-CE 14	AGGTGTTTCGTGAAGAGCAGG	GAACCGACAACCAAAAGCCC	(AGG) ₅	220	Sim
RNA-CE 15	TTCTTTCTGACCGACCGTG	GCCACCATCACCACCAAAAC	(ATCC) ₈	117	Não
RNA-CE 16	ATCTCCTCCTCCACCTCACC	CATCAGCTTGGACCTACGCA	(GCG) ₆	294	Não
RNA-CE 17	TGGTGGTGCTTTGTTCAAGT	GCTTCTCCAAACGCCACATC	(AGA) ₉	271	Sim
RNA-CE 18	TGGATGATCCACGGTGCAAA	ATTGTAGCAAAGCCCCGCTT	(TTG) ₅	371	Sim
RNA-CE 19	ACTAGTCACACACACAGGCG	CCCACCATGGCTTGTCTTCT	(GGAT) ₅	195	Sim

RNA-CE 20	GATGACGACGACGATGACGA	TACCCCTCCAGCTTCTCCAG	(CGA) ₅	146	Não
RNA-CE 21	CCGTGTTGAATTGCTCCGTG	ATGTTCTTGGAGAGGCAGGC	(GCG) ₅	143	Não
RNA-CE 22	AAAGAGGAGAGGGGCTAGGG	TGTTGGTGGCCTGGTCAAAT	(GGC) ₅	256	Sim
RNA-CE 23	CCCTCATCTCCACGCTCAAG	GGATGAGGAGGCTGAGGTTG	(CGT) ₅	158	Sim
RNA-CE 24	ACGATCAAGGACAAGTCGCC	GCCTCTAGTTCTCGAAGGCC	(GCA) ₅	121	Não
RNA-CE 25	TCCTCCCTCTCTGTGCTC	TACCCCTGTCCGATCTTCGT	(GCG) ₅	309	Sim
RNA-CE 26	CTGCAGAGCTCCACAGAACA	CCTGCAGGATCGTGTAGTCC	(GCC) ₅	124	Não
RNA-CE 27	TCACAGGAGGAGACCGATGT	CCTGTCTGCGAAGTTCACCT	(CAC) ₅	363	Não
RNA-CE 28	CTCTCTCCTCCATCCTCCCC	GGGGAAGGAGGAGAGGATGT	(TCT) ₁₂	217	Não
RNA-CE 29	CAGCCAGGTCATCCTCAGTC	TTTAGCAAACAAGCCGCCG	(GGC) ₅	362	Não
RNA-CE 32	GTCGGGGTCGTTCAAGAAGT	GCATCCACGTCCTCGAAGAA	(GGC) ₅	132	Não
RNA-CE 33	AGGCGCAAGGGATGAATGAA	CCTATCTCGCCGTCTCACAC	(GGC) ₅	291	Não
RNA-CE 34	CTTCCCCTCATCACACCACC	GGCTGAGAAGAGGGTGTTC	(GCT) ₆	261	Sim
RNA-CE 35	CTTCTCCTTCGCCTCATCCC	TAAGAAGGGGATGAGGCGGA	(CCT) ₅	121	Não
RNA-CE 36	TGAGTCCCAAGAAGCAGCAG	TGCTGTTTGGCTCGATCCAT	(CGG) ₆	356	Sim
RNA-CE 37	TTAATGCCGCTGCGATGTTG	CATCTAGCCACAGGTGCACA	(TAG) ₅	389	Não
RNA-CE 38	CTAGCTTTGCTTGCCACTGC	GCACAGCAGACATGGATCCT	(CA) ₆	243	Não
RNA-CE 39	ATCACAGCAAGAGGAGCCAC	TGTAGTGCTCGGGATCCTCA	(CAC) ₅	345	Sim
RNA-CE 40	CCGCAAATCCTCAGAACCCT	GTGCGGTGGATTTTGCTTGT	(CCG) ₅	267	Não
RNA-CE 41	AGACCCCTACACGAGCTTCT	CCGGGTACTGATGATGAGGC	(CGC) ₇	164	Não
RNA-CE 42	GCTGCTCTGTCTCCACTTGT	GCTTACGAGGTTCCGGAACA	(GCG) ₅	276	Sim
RNA-CE 43	AATACTCTCCCCTCCCCAC	CGCTCTCGTACTACCAGCAG	(AG) ₈	149	Sim
RNA-CE 44	GTGCGAGAGGGAAACACAGA	TCGGTGTGCTTGTAGTGGAC	(GCG) ₅	155	Sim
RNA-CE 45	TCCTAGCTGACCGGACTACC	AGGCTTTAGCAACCGAAGCT	(CGG) ₆	365	Não

RNA-CE 46	GAGAGCGAGAGACATGAGGC	ACAGGCCAAGCAAGAGGTAC	(GGC) ₇	137	Sim
RNA-CE 47	TGCCGAGGACAGAAGAAGTG	CAGACGTGCTCATCACCTCA	(AGG) ₅	381	Não
RNA-CE 48	TATACATGCCAGCGACGAC	TCCTAGCCTTCTCTGCGTCT	(GT) ₈	383	Sim
RNA-CE 49	GACATCCTCGTCGTCGTCTC	CCTAGTTCTTACCGGTGGCG	(CGA) ₅	232	Sim
RNA-CE 50	AAGGGGAAGAAGTGCTACGC	GATCACCCAAGAACGTTGCG	(GGC) ₅	117	Não

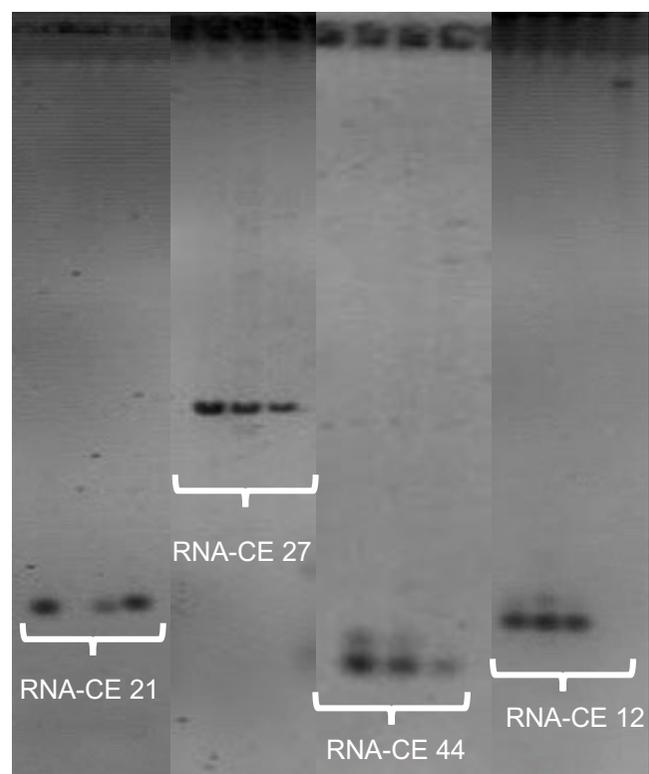


Figura 13: Géis de agarose com os *primers* RNA-CE 21, RNA-CE 27, RNA-CE 44 e RNA-CE 12 nas amostras BAGCE 11, BAGCE50, BAGCE 57 e BAGCE 69 de *C. purpureus* inicialmente testadas.

3.3.2 PAINEL ESPECÍFICO PARA IDENTIFICAÇÃO DE CULTIVARES DE *Cenchrus purpureus*

A partir dos resultados dos testes de amplificação dos *primers* desenhados, os 21 *primers* promissores foram avaliados para a identificação de polimorfismos entre as quatro cultivares (BRS Kurumi, BRS Canará, Pioneiro e BRS Capiáçu) já lançadas pelo Programa de Melhoramento de capim elefante da Embrapa e acessos pertencentes ao BAGCE. Entre os acessos avaliados do BAGCE estavam os progenitores da cultivar BRS Capiáçu, que são os acessos BAG 60 e BAG 57.

A amplificação do primer RNA-CE 05 (Figura 14) permitiu a identificação de um padrão de bandas único para cultivar BRS Kurumi. Apesar de compartilhar as bandas 275pb e 295pb com o acesso BAG 58, ainda é possível diferenciar as duas amostras pela presença de uma banda a mais (280pb) na cultivar BRS Kurumi.

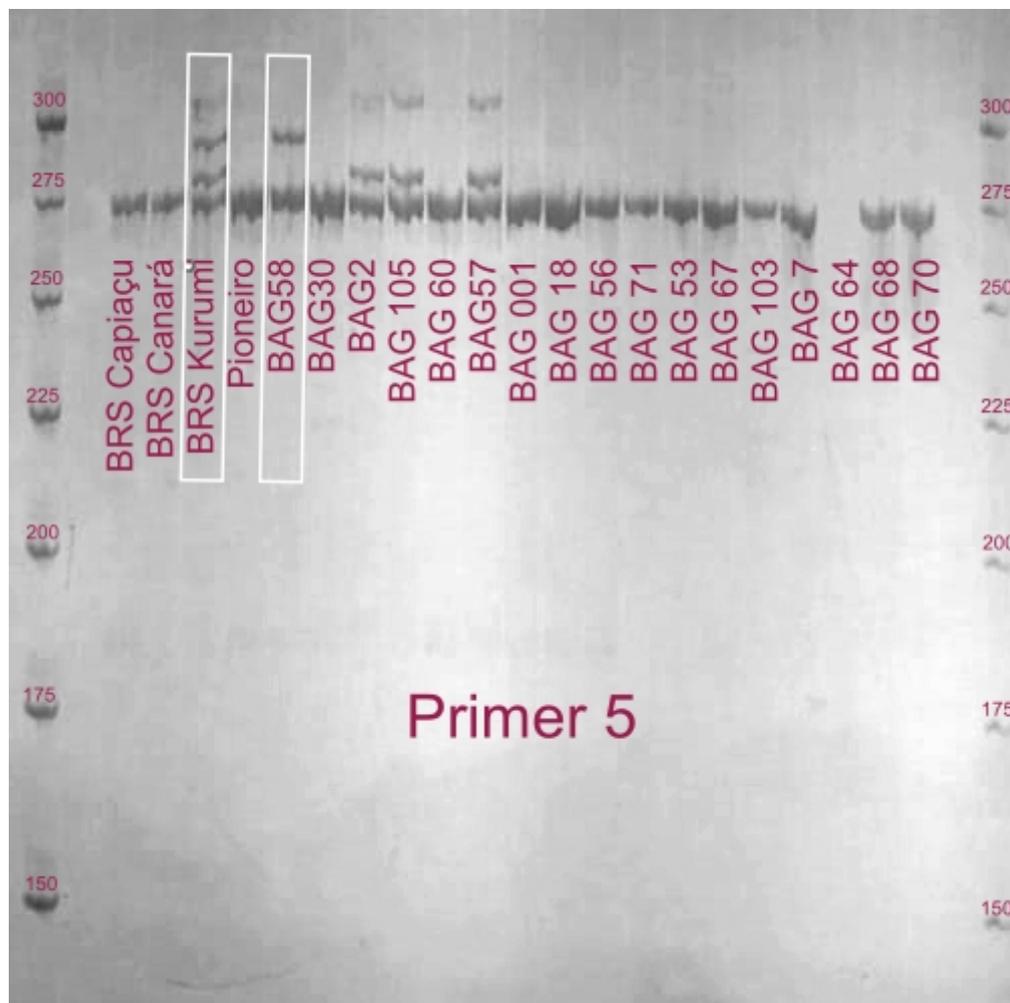


Figura 14: Gel de poliacrilamida com o *primer* RNA-CE 05 nas 21 cultivares/aceessos de capim-elefante avaliadas. A primeira e última columnas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado.

Foi possível observar, utilizando o *primer* RNA-CE 44, um padrão de bandas único para BRS Canará, como visto na Figura 15. Para esta cultivar, foi identificado um perfil de bandas nos tamanhos de 145, 150 e 158 pb.

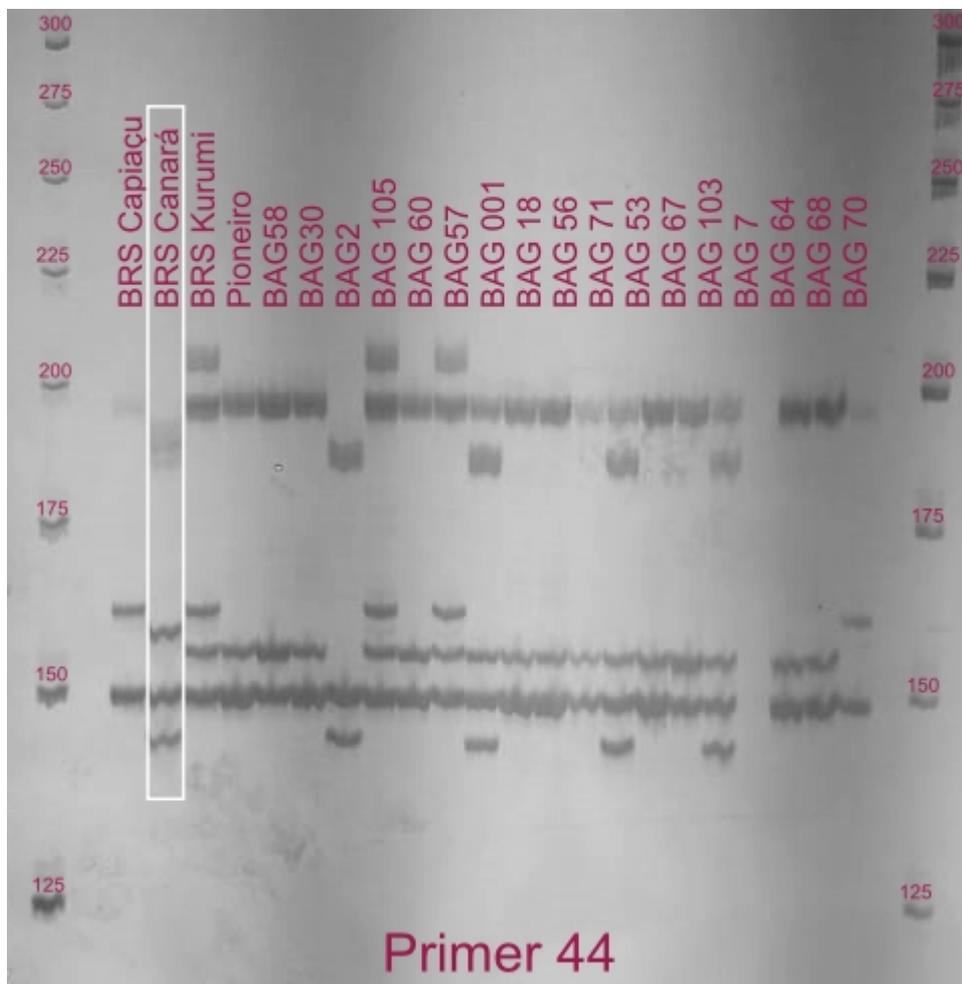


Figura 15: Gel de poliacrilamida com o *primer* RNA-CE 44 nas 21 cultivares/aceessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado.

A presença do fragmento de 133 pb na cultivar Pioneiro (Figura 16), utilizando o *primer* RNA-CE 46, pode ser considerada uma banda rara e importante na identificação desta cultivar. Apesar desta banda aparecer nos acessos BAG 1 e BAG 7, foi possível identificar um padrão de bandas diferente

para Pioneiro, se comparado às cultivares BRS Capiaçú, BRS Canará e BRS Kurumi.

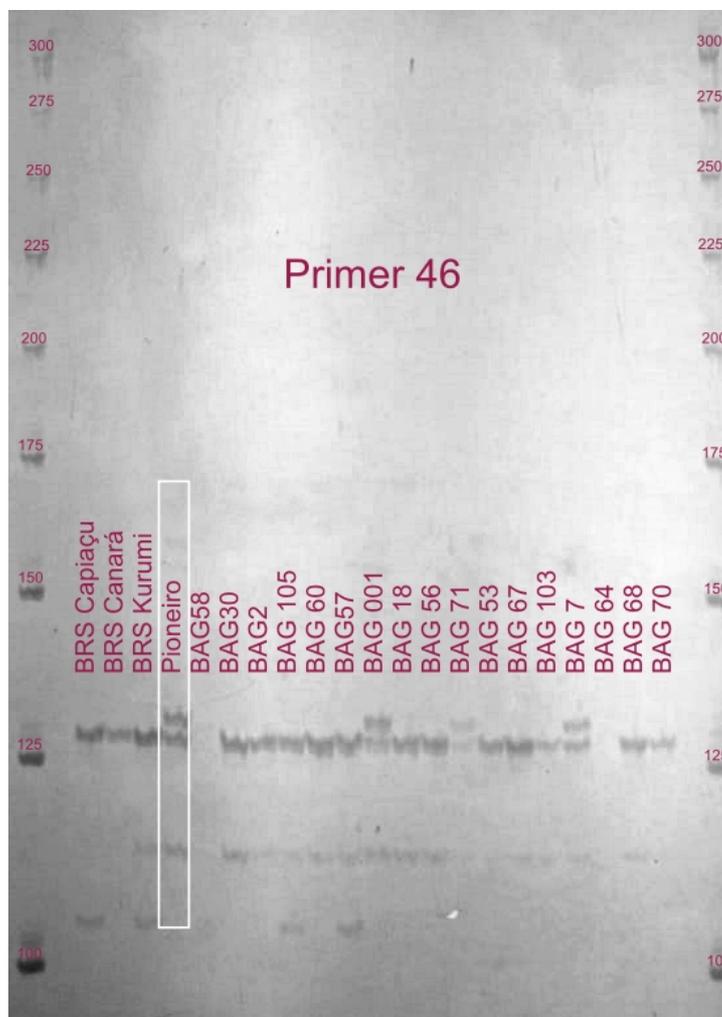


Figura 16: Gel de poliacrilamida com o *primer* RNA-CE 46 nas 21 cultivares/aceessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases de cada fragmento amplificado.

Para BRS Capiaçú, três *primers* em conjunto foram informativos para a diferenciação desta cultivar. O *primer* RNA-CE 17 amplificou um fragmento raro de 270 pb (Figura 17). O RNA-CE 43 permitiu a diferenciação da cultivar pela ausência de bandas (Figura 18). Esse primer foi avaliado em diferentes condições e com várias amostras da cultivar BRS Capiaçú, em todos os testes não foi possível amplificar qualquer fragmento.

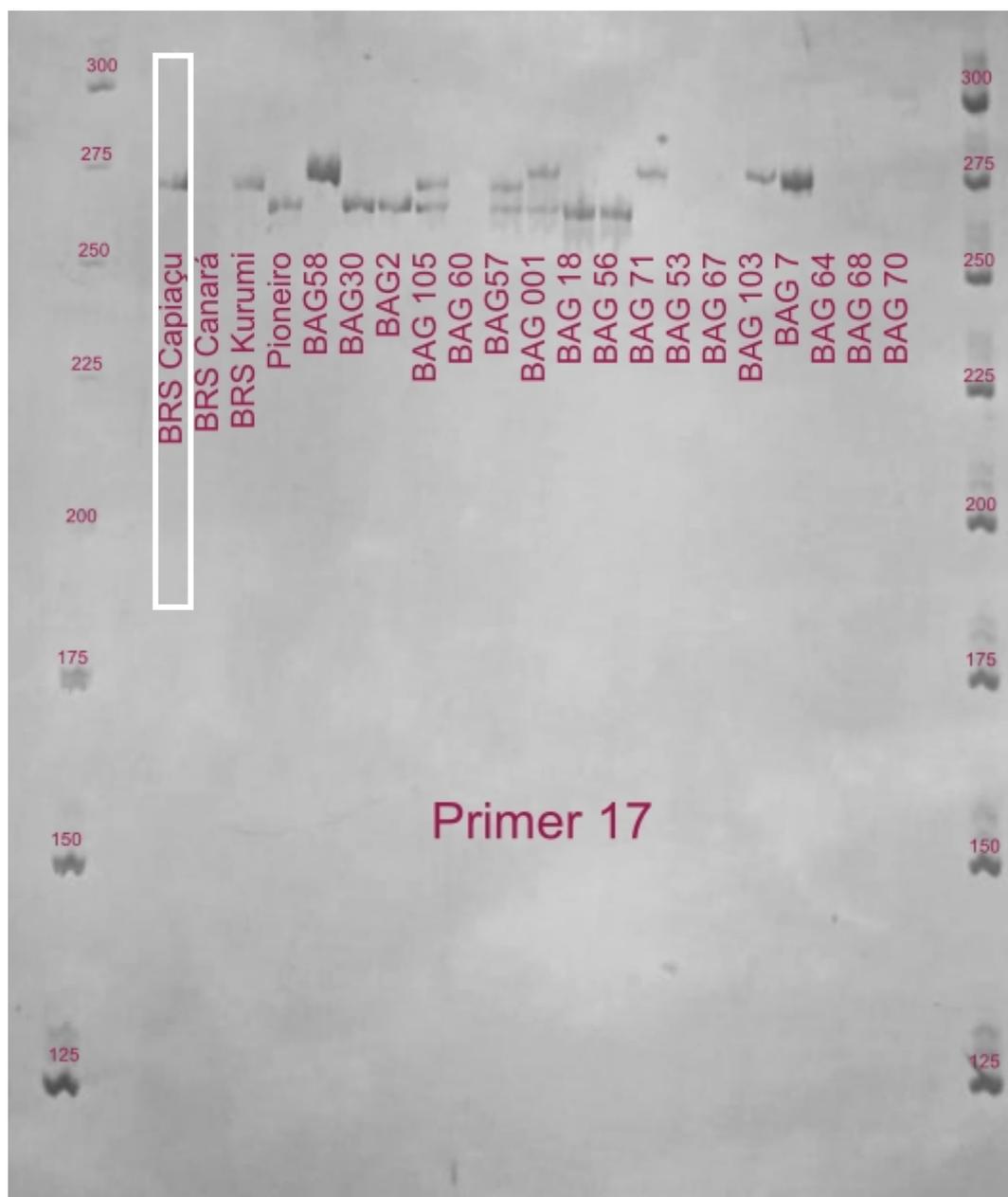


Figura 17: Gel de poliacrilamida com o *primer* RNA-CE 17 nas 21 cultivares/aceessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases em cada fragmento amplificado.

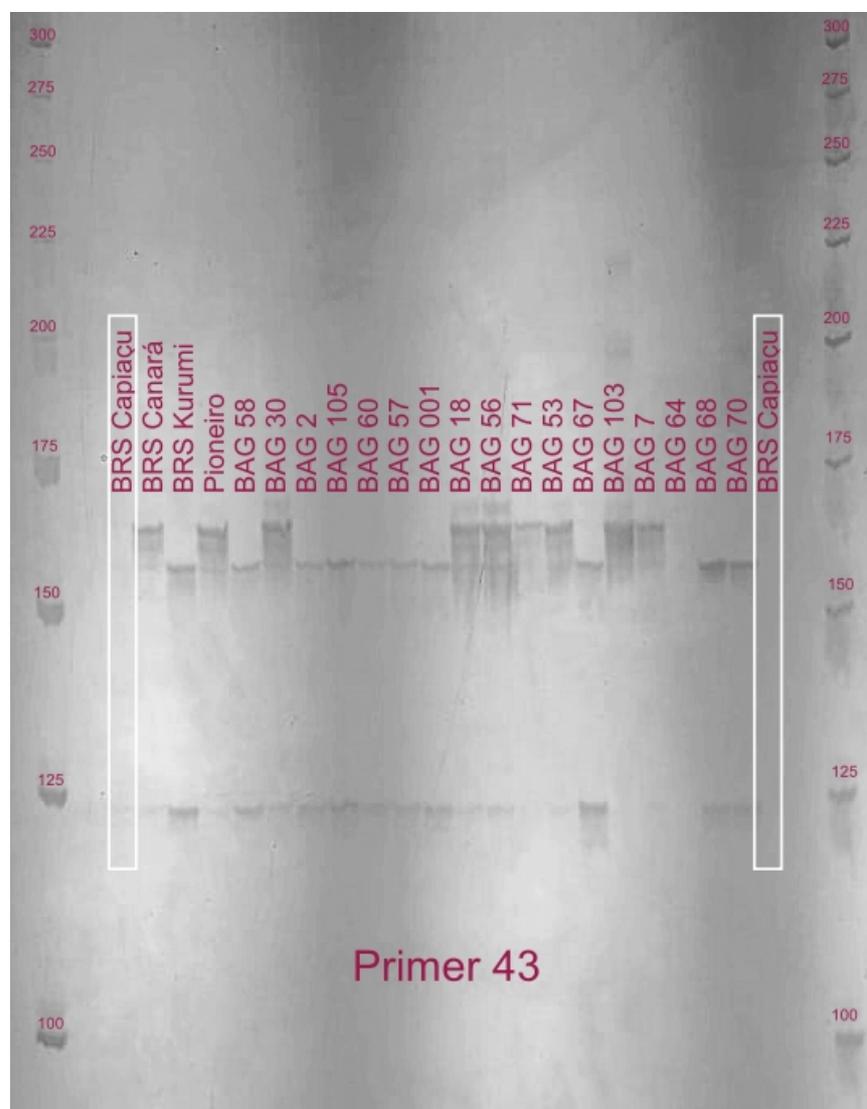


Figura 18: Gel de poliacrilamida com o *primer* RNA-CE 43 nas 21 cultivares/ acessos de capim-elefante avaliadas. A primeira e última colunas são padrões de pesos moleculares usados para inferir os tamanhos em pares de bases em cada fragmento amplificado.

Tabela 12: Tamanho de bandas amplificadas nas 21 cultivares/acessos de capim-elefante, considerando os *primers* mais promissores na identificação de um padrão único nas cultivares BRS Capiaçú, BRS Canará, BRS Kurumi e Pioneiro.

Cultivares ou Acessos	Primers				
	RNA-CE 05	RNA-CE 17	RNA-CE 43	RNA-CE 44	RNA-CE 46
BRS Capiaçú	275	265	-	150/162	107/128
BRS Canará	275	290/300	120/165	146/150/160	128
BRS Kurumi	275/280/295	-	155	150/158/162/198/210	107/117/128
Pioneiro	275	260/300	165	150/158/198	117/128/130
BAG 58	275/295	270/290	120/155	150/158/198	-
BAG 30	275	260/300	165	150/158/198	117/128
BAG 2	275/280	260/290/300	120/155	146/150/190	128
BAG 105	275/280	260/270/290/>300	120/155	150/158/160/162/198/210	107/128
BAG 60	275	290/300	120/155	150/158/198	117/128
BAG 57	275/280	270/>300	120/155	150/158/162/198/210	107/117/128
BAG 1	275	270/>300	120/155	146/150/158/190/198	117/128/130
BAG 18	275	260/300	155/165	150/158/198	117/128
BAG 56	275	260/300	120/155/165	150/158/198	117/128
BAG 71	275	270	165	150/158	-
BAG 53	275	290/300	165	146/150/158/190	128
BAG 67	275	280/290	120/155	150/158/198	128
BAG 103	275	270/300	165	150/158/198	128
BAG 7	275	270	165	146/150/158/190/198	117/128/130
BAG 64	-	-	-	-	-
BAG 68	275	280/290	120/155	150/158/198	117/128
BAG 70	275	280/290	120/155	150/158/198	128

Na avaliação dos *primers*, foram utilizadas 21 cultivares/acessos para demonstrar a ampla diversidade encontrada na espécie. Além disso, uma maior amostragem das cultivares de capim-elefante ratificaram a existência de um padrão de bandas único nas cultivares BRS Kurumi, BRS Canará, BRS Capiáçu e Pioneiro, importantes no mercado de forrageiras. Esta metodologia foi também utilizada em estudos com maçã (BARIC *et al.*, 2020) e cana-de-açúcar (SINGH *et al.*, 2019) com o objetivo de identificar *fingerprints* de cultivares importantes. A caracterização das cultivares BRS Capiáçu, BRS Kurumi, BRS Canará e Pioneiro a nível molecular poderá ser uma ferramenta importante na confirmação da identidade de plantas no mercado de forrageiras, protegendo a propriedade intelectual da cultivar.

3.4 CONCLUSÃO

Os *primers* desenhados a partir do transcriptoma das espécies foram eficazes na amplificação em amostras testadas de *U. humidicola* e *C. purpureus*. Na identificação de padrão único nas cultivares BRS Kurumi, BRS Canará, BRS Capiáçu e Pioneiro, destacaram-se 5 *primers*, que poderão ser usados como *fingerprint* do DNA na garantia de pureza das cultivares no mercado de forrageiras.

REFERÊNCIAS BIBLIOGRÁFICAS

ABIEC. Associação Brasileira das Indústrias Exportadoras de Carne. Perfil da Pecuária no Brasil 2019. Disponível em: <http://www.abiec.com.br/control/uploads/arquivos/sumario2019portugues.pdf>. Acesso em 28 maio 2019.

ABRIL, J.F.; CASTELLANO, S. Genome Annotation. In: RANGANATHAN, S.; GRIBSKOV, M.; NAKAI, K.; SCHÖNBACH, C. (Org.). Encyclopedia of Bioinformatics and Computational Biology, Academic Press, v. 2, p. 195-209, 2019.

ADAMOWSKI, E.V.; BOLDRINI, K.R.; PAGLIARINI, M.S.; DO VALLE, C.B.; Abnormal cytokinesis in microsporogenesis of *Brachiaria humidicola* (Poaceae: Paniceae). Genetics and Molecular Research, v.6, p. 616-621, 2007.

ADAMS, S.M.; LINDMEIER, J.B.; DUVERNELL, D.D. Microsatellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, *Fundulus heteroclitus*. Molecular Ecology, v. 15, p. 1109-1123, 2006.

ADESOGAN, A.T.; VENDRAMINI, J.; SOLLENBERGER, L.; NEWMAN, Y. Defining Forage Quality. EDIS, 2017.

ALCÂNTARA, P.B.; BUFARAH, G. Plantas forrageiras: gramíneas e leguminosas. São Paulo: Editora Nobel, 1983.

ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. Journal of Molecular Biology, v. 215, p. 403-410, 1990.

ANEEL. Biomassa. In: AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (BRASIL). Atlas de energia elétrica no Brasil / Agência Nacional de Energia Elétrica – Brasília:ANEEL, 2005.

ANDERSON, W.F.; DIEN, B.S.; BRANDON, S.K.; PETERSON, J. Assessment of bermudagrass and bunch grasses as feedstocks for conversion to ethanol. Applied Biochemistry and Biotechnology, v. 145, p. 13–21, 2008.

ANDRADE, C.M.S.; ASSIS, G.M.L. Capim-Xaraés: Cultivar de Gramínea Forrageira Recomendada para Pastagens no Acre. Rio Branco, AC: Embrapa Acre, 2008.

ARRUDA, P. Genetically modified sugarcane for bioenergy generation. *Current Opinion in Biotechnology*, v. 23, p. 315-322, 2012.

AZEVEDO, A.L.S.; COSTA, P.P.; MACHADO, J.C.; MACHADO, M.A.; PEREIRA, A.V.; LÉDO, F.J.S. Cross species amplification of *Pennisetum glaucum* microsatellite markers in *Pennisetum purpureum* and genetic diversity of Napier grass accessions. *Crop Science*, v. 52, p. 1776-1785, 2012.

BAJWA, D.S.; POURHASHEM, G.; ULLAHM A.H.; BAJWA, S.G. A concise review of current lignin production, applications, products and their environmental impact. *Industrial Crops and Products*, v. 139, p. 111536, 2019.

BALL, D.M.; COLLINS, M.; LACEFIELD, G.D.; MARTIN, N.P.; MERTENS, D.A.; OLSON, K.E.; PUTNAM, D.H.; UNDERSANDER, D.J.; WOLF, M.W. Understanding Forage Quality. American Farm Bureau Federation Publication 1-01, Park Ridge, IL, 2001.

BALMANT, K.M.; NOBLE, J.D.; ALVES, F.C.; DERVINIS, C.; CONDE, D.; SCHMIDT, H.W.; VAZQUEZ, A.I.; BARBAZUK, W.B.; DE LOS CAMPOS, G.; RESENDE JUNIOR, M.F.R.; KIRST, M. Xylem systems genetics analysis reveals a key regulator of lignin biosynthesis in *Populus deltoides*. *Genome research*, v. 30, n.8, p. 1131-1143, 2020.

BAMDAD, H.; HAWBOLDT, K.; MacQUARRIE, S. A review on common adsorbents for acid gases removal: Focus on biochar. *Renewable and Sustainable Energy Reviews*, 2017.

BANG, K.H.; JO, I.H.; CHUNG, J.W.; KIM, Y.C.; LEE, J.W.; SEO, A.Y.; PARK, J.H.; KIM, O.T.; HYUN, D.Y.; KIM, D.H.; CHA, S.W. Analysis of Genetic Polymorphism of Korean Ginseng Cultivars and Foreign Accessions Using SSR Markers. *Korean Journal of Medicinal Crop Science*, v. 19, p. 347-353, 2011.

BARIC, S.; STORTI, A.; HOFER, M.; GUERRA, W.; VIA, J.D. Molecular Genetic Identification of Apple Cultivars Based on Microsatellite DNA Analysis. I. The Database of 600 Validated Profiles. *Erwerbs-Obstbau*, v. 62, n. 2, p. 117-154, 2020.

BARROS, J.; SERK, H.; GRANLUND, I.; PESQUET, E. The cell biology of lignification in higher plants. *Annals of Botany*, v. 115, n. 7, p. 390-395, 2015.

BARROS, J.; SERRANI-YARCE, J.C.; CHEN, F.; BAXTER, D.; VENABLES, B.J.; DIXON, R.A. Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nature Plants*, 2016.

BAUCHER, M.; HALPIN, C.; PETIT-CONIL, M.; BOERJAN, W. Lignin: Genetic engineering and impact on pulping. *Critical Reviews in Biochemistry and Molecular Biology*, v. 38, p. 305-350, 2003.

BEIER, S.; THIEL, T.; MÜNCH, T.; SCHOLZ, U.; MASCHER, M. MISA-web: a web server for microsatellite prediction. *Bioinformatics*, v. 33, p. 2583-2585, 2017.

BELALIA, N.; LUPINI, A.; DJEMEL, A.; MORSLI, A.; MAUCERI, A.; LOTTI, C.; KHELIFI-SLAOUI, M.; KHELIFI, L.; SUNSERI, F. Analysis of genetic diversity and population structure in Saharan maize (*Zea mays* L.) populations using phenotypic traits and SSR markers. *Genetic Resources and Crop Evolution*, v. 66, p. 243-257, 2019.

BHATIA, R.; GALLAGHER, J.A.; GOMEZ, L.D.; BOSCH, M. Genetic engineering of grass cell wall polysaccharides for biorefining. *Plant Biotechnology Journal*, v. 15, n. 9, p. 1071-1092, 2017.

BIDLACK, J.E.; DASHEK, W.V. *Plant Cells and their Organelles*, John Wiley & Sons, Ltd, 2017, p. 209-238.

BISWAL, A.K.; HAO, Z.Y.; PATTATHIL, S.; YANG, X.H.; WINKELER, K.; COLLINS, C.; MOHANTY, S.S.; RICHARDSON, E.A.; GELINEO-ALBERSHEIN, I.; HUNT, K.; RYNO, D.; SYKES, R.W.; TURNER, G.B.; ZIEBELL, A.; GJERSING, E.; LUKOWITZ, W.; DAVIS, M.F.; DECKER, S.R.; HAHN, M.G.; MOHNEN, D. Downregulation of GAUT12 in *Populus deltoides* by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. *Biotechnology for Biofuels*, v. 8, n. 1, p. 1-26, 2015.

BLAVET, N.; CHARIF, D.; OGER-DESFEUX, C.; MARAIS, G.A.B.; WIDMER, A. Comparative high-throughput transcriptome sequencing and development of SiESTa, the *Silene* EST annotation database. *BMG Genomics*, v. 12, p. 376, 2011.

BOERJAN, W.; RALPH, J. E BAUCHER, M. Lignin biosynthesis. *Annual Review of Plant Biology*, v. 54, p. 519-546, 2003.

BOGDAN, A. (1977) *Tropical pastures and fodder plants*. Longman, London.

BOIJA, E.; JOHANSSON, G. Interactions between model membranes and lignin-related compounds studied by immobilized liposome chromatography. *Biochimica et Biophysica Acta (BBA) – Biomembranes*, v. 1758, n. 5, p. 620-626, 2006.

BOLDRINI, K.R.; PAGLIARINI, M.S.; DO VALLE, C.B. Meiotic behavior of a nonaploid accession endorses $x=6$ for *Brachiaria humidicola* (Poaceae). *Genetics and Molecular Research*, v. 8, p. 1444-1450, 2009a.

BOLDRINI, K.R.; MICHELETTI, P.L.; GALLO, P.H.; MENDES-BONATO, A.B.; PAGLIARINI, M.S.; DO VALLE, C.B. Origin of a polyploid accession of *Brachiaria humidicola* (Poaceae: Panicoideae: Paniceae). *Genetics and Molecular Research*, v. 8, p. 888-895, 2009b.

BOLDRINI, K.R.; PAGLIARINI, M.S.; DO VALLE, C.B. Evidence of natural hybridization in *Brachiaria humidicola* (Rendle) Schweick. (Poaceae: Panicoideae: Paniceae). *Journal of Genetics*, v. 89, p. 91-94, 2010.

BOLGER, A.M.; LOHSE, M.; USADEL, B. Trimmomatic: A flexible read trimming tool for Illumina NGS data. *Bioinformatics*, 2014.

BONAWITZ, N.D.; CHAPPLE, C. The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. *Annual Review of Genetics*, v. 44, p. 337-363, 2010.

BONCAN, D.A.T.; TSANG, S.S.K.; LI, C.; LEE, I.H.T.; LAM, H.M.; CHAN, T.F.; HUI, J.H.I. Terpenes and terpenoids in Plants: Interactions with Environment and Insects. *International Journal of Molecular Sciences*, v. 21, p. 7382, 2020.

BORA, A.; CHOUDHURY, P.R.; PANDE, V.; MANDAL, A.B. Assessment of genetic purity in rice (*Oryza sativa*L.) hybrids using microsatellite markers. *3 Biotech*, v. 6, n. 1, p. 50, 2016.

BORIN, G.P.; SANCHEZ, C.C.; DE SANTANA, E.S.; ZANINI, G.K.; DOS SANTOS, R.A.C.; PONTES, A.O.; DE SOUZA, A.T.; DAL'MAS, R.M.M.T.S.; RIAÑO-PACHÓN, D.M.; GOLDMAN, G.H.; OLIVEIRA, J.V.C. Comparative transcriptome analysis reveals different strategies for degradation of steam-exploded sugarcane bagasse by *Aspergillus niger* and *Trichoderma reesei*. *BMC Genomics*, v. 18, p. 501, 2017.

BOURGON, R.; GENTLEMAN, R.; HUBER, W. Independent filtering increases detection power for high-throughput experiments. *Proceedings of the National Academy of Sciences*, v. 107, p. 9456-9551, 2010.

BRITO, C.J.F.A.; RODELLA, R.A. E DESCHAMPS, F.C. Perfil Químico da Parede Celular e suas Implicações na Digestibilidade de *Brachiaria brizantha* e *Brachiaria humidicola*. *Revista Brasileira de Zootecnia*, v. 32, n.6, p. 1835-1844, 2003.

BROWN, D.M.; GOUBERT, F.; WONG, V.W.; GOODACRE, R.; STEPHENS, E.; DUPREE, P.; TURNER, S.R. Comparison of five xylan synthesis mutants reveals new insight into the mechanism of xylan synthesis. *Plant Journal*, v. 52, p. 1154-1168, 2007.

BRUNKEN J.N. A systematic study of *Pennisetum* sect. *Pennisetum* (Gramineae). *American Journal of Botany*, v. 64, p. 161–176, 1977.

BRYANT, D.M.; JOHNSON, K.; DI TOMASSO, T.; TICKLE, T.; COUGER, M.B.; PAYZIN-DOGRU, D.; LEE, T.J.; LEIGH, N.D.; KUO, T.H.; DAVIS, F.G.; BATEMAN, J.; BRYANT, S.; GUZIKOWSKI, A.R.; TSAI, S.L.; COYNE, S.; YE, W.W.; FREEMAN JUNIOR, R.M.; PESHKIN, L.; TABIN, C.J.; REGEV, A.; HAAS, B.J.; WHITED, J.L. A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors. *Cell Reports*, v. 18, n. 3, p. 762-776, 2017.

CAIXETA, P.V. SIMEONE, M.L.F.; GONTIJO, M.M.; NETO, GUIMARÃES, C.C.; GOMES, P.C.; MEDEIROS, E.P. Desenvolvimento de um modelo de regressão multivariado para determinação do teor de matéria seca em braquiária empregando espectroscopia no infravermelho próximo (NIRS). In: Seminário de iniciação científica. Sete Lagoas, MG: Embrapa Milho e Sorgo (CNPMS), p. 5, 2013.

CANILHA, L.; CHANDEL, A.K.; MILESSI, T.S.S.; ANTUNES, F.A.F.; FREITAS, W.L.C., FELIPE, M.G.A. DA SILVA, S.S. Bioconversion of sugarcane biomasses into ethanol: an overview about composition, pretreatment methods, detoxification of hydrolysates, enzymatic saccharification, and ethanol fermentation. *Journal of Biomedicine and Biotechnology*, 2012.

CARDOSO-SILVA, C.B.; COSTA, E.A.; MANCINI, M.C.; BALSALOBRE, T.W.A.; CANESIN, L.E.C.; PINTO, L.R.; CARNEIRO, M.S.; GARCIA, A.A.F.; DE SOUZA, A.P.; VICENTINI, R. *De novo* Assembly and Transcriptome Analysis of Contrasting Sugarcane Varieties. *PLoS ONE*, v. 9, n. 2, 2014.

CARNEIRO, A.C.G.; NUNEZ, H.; MORAES, M.M.G.A.; ONAL, H. An Economic Analysis of Land Use Changes and Biofuel Feedstock Production in Brazil: The Role of Irrigation Water. Disponível em:

<http://www.webmeets.com/files/papers/wcere/2014/993/WCERE%202014%20Onal%20et%20al.pdf>. Acesso em 31 outubro 2017.

CAROCHA, V. SOLER, M.; HEFER, C.; CASSAN-WANG, H.; FEVEREIRO, P.; MYBURG, A.A.; PAIVA, J.A.P.; GRIMA-PETTENATI, J. Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytologist*, 2015.

CASTILLA, C.; JACKSON, W.A. Nitrogen uptake mechanism in two *Brachiaria* species: Y-315. In: TropSoils technical report, 1988-1989. TropSoils, Raleigh, 1991, p. 194-167.

CASTILLO, R.M.; FENG, K.; HUBACEK, K.; SUN, L.; GUILHOTO, J.; MIRALLES-WILHELM, F. Uncovering the Green, Blue, and Grey Water Footprint and Virtual Water of Biofuel Production in Brazil: A Nexus Perspective. *Sustainability*, v. 9, p. 2049, 2017.

CAVALCANTE, M.; LIRA, M.A. Variabilidade Genética em *Pennisetum purpureum* Shumacher. *Revista Caatinga*, v. 23, n. 2, p. 153-163, 2010.

CESARINO, I.; SIMÕES, M.S.; BRITO, M.S.; FANELLI, A.; SILVA, T.F.; ROMANEL, E. Building the wall: recente advances in understanding lignina metabolism in grasses. *Acta Physiologiae Plantarum*, v. 38, p 269, 2016.

CHEMISQUY, M.A.; GIUSSANI, L.M.; SCATAGLINI, M.A.; KELLOGG, E.A.; MORRONE, O. Phylogenetic studies favour the unification of *Pennisetum*, *Cenchrus* and *Odontelytrum* (Poaceae): a combined nuclear, plastid and morphological analysis, and nomenclatural combinations in *Cenchrus*. *Annals of Botany*, v. 106, p.107-130, 2010.

CHEN, F.; DIXON, R.A. Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology*, Nature Publishing Group, v. 25, p. 759-761, 2007.

CHEN, Z.; TANG, D.; NI, J.; LI, P.; ZHOU, J.; LI, C.; LAN, H.; LI, L.; LIU, J. Development of genic KASP SNP markers from RNA-Seq data for map-based cloning and marker-assisted selection in maize. *BMC Plant Biology*, v. 21, n. 1, p. 157, 2021.

CHERNEY, J.H.; CHERNET, D.J.R. AKIN, D.E.; AXTELL, J.D. Potential of brown-midrib, low-lignin mutants for improving forage quality. *Advances in Agronomy*, v. 46, p. 157-195, 1991.

CHERNEY, J.H.; HALL, M.H. Forage Quality in Perspective. Penn State Extension, 2008.

CHOI, J.K.; SA, K.J.; PARK, D.H.; LIM, S.E.; RYU, S.H.; PARK, J.Y.; PARK, K.J.; RHEE, H.I.; LEE, M.; LEE, J.K. Construction of genetic linkage map and identification of QTLs related to agronomic traits on DH population of maize (*Zea mays* L.) using SSR markers. *Genes & Genomics*, v. 41, p. 667-678, 2019.

CHRISTENSEN, J.H.; BAUCHER, M.; O'CONNELL, A.P.; VAN MONTAGY, M.; BOERJAN, W. Control of lignin biosynthesis. In: JAIN, S.M. E MINOCHA, S.C. *Molecular*

Biology of Wood Plants, Volume 1, For. Sci., v. 64, p. 227-267. Dordrecht: Kluwer. 2000, 520 p.

CLOONAN, N.; FORREST, A.R.R.; KOLLE, G.; GARDINER, B.B.A.; FAULKNER, G.J.; BROWN, M.K.; TAYLOR, D.F.; STEPTOE, A.L.; WANI, S.; BETHEL, G.; ROBERTSON, A.J.; PERKINS, A.C.; BRUCE, S.J.; LEE, C.C.; RANADE, S.S.; PECKHAM, H.E.; MANNING, J.M.; MCKERNAN, K.J.; GRIMMOND, S.M. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods*, v. 5, p. 613-619, 2008.

CLOUSE, S.D. Arabidopsis mutants reveal multiples roles for sterols in plant development. *Plant Cell*, v. 14, p. 1995-2000, 2002.

COSGROVE, D.J.; JARVIS, M.C. Comparative structure and biomechanics of plant primary and secondary cell walls. *Frontiers in Plant Sciences*, v. 3, p. 204, 2012.

DAVIDE, L.C.; TECHIO, V.H.; NUNES, J.D.; PEREIRA, A.V. Variação cromossômica numérica em *Pennisetum*. *Ciência e Agrotecnologia*, v. 31, n. 2, p. 398-405, 2007.

DA SILVA, I.S.; Transcriptoma da via de biossíntese dos monolignóis ao longo do desenvolvimento dos entrenós de cana-de-açúcar (*Saccharum* spp.). 2019. 69 f. Dissertação (Mestrado em Biotecnologia Industrial) – Universidade de São Paulo, Lorena.

DEL RÍO, J.C.; PRINSEN, P.; RENCORET, J.; NIETO, L.; JIMÉNEZ-BARBERO, J.; RALPH, J.; MARTÍNEZ, Á.T.; GUTIÉRREZ, A. Structural characterization of the lignin in the cortex and pith of elephant grass (*Pennisetum purpureum*) stems. *Journal of Agricultural and Food Chemistry*, v. 60, n. 14, p. 3619-3634, 2012a.

DEL RÍO, J.C.; RENCORET, J.; PRINSEN, P.; MARTÍNEZ, Á.T.; RALPH, J.; GUTIÉRREZ, A. Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR and reductive cleavage methods. *Journal of Agricultural and Food Chemistry*, v. 60, n. 23, p. 5922-5935, 2012b.

DELLA-BIANCA, B.E.; BASSO, T.O.; STAMBUK, B.U.; BASSO, L.C.; GOMBERT, A.K. What do we know about the yeast strains from the Brazilian fuel ethanol industry? *Applied Microbiology and Biotechnology*, v. 97, p. 979-991, 2013.

DENYER, T.; MA, X.; KLESEN, S.; SCACCHI, E.; NIESELT, K.; TIMMERMANS, M.C.P. Spatiotemporal development trajectories in the Arabidopsis root revealed using high-throughput single-cell RNA sequencing. *Developmental Cell*, v. 48, n. 6, p. 840-852, 2019.

DERESZ, F. Capim-elefante manejado em sistema rotativo para produção de leite e carne. In: PASSOS, L.P.; CARVALHO, L.A.; MARTINS, C.E.; BRESSAN, M.; PEREIRA, A.V. (Ed.). *Biologia e manejo do capim-elefante*. Juiz de Fora: Embrapa Gado de Leite, 1999a. p. 161-172.

DERESZ, F. 1999b. Utilização do capim-elefante sob pastejo rotativo para produção de leite e carne. Juiz de Fora, Embrapa-CNPGL, 29p. (Circular técnica 54).

DJEBALI, S.; DAVIS, C.A.; MERKEL, A.; DOBIN, A.; LASSMANN, T.; MORTAZAVI, A.; TANZER, A.; LAGARDE, J.; LIN, W.; SCHLESINGER, F.; XUE, C.; MARINOV, G.K.; KHATUN, J.; WILLIAMS, B.A.; ZALESKI, C.; ROZOWSKY, J.; RÖDER, M.; KOKOCINSKI, F.; ABDELHAMID, R.F.; ALIOTO, T.; ANTOSHECHKIN, I.; BAER, M.T.; BAR, N.S.; BATUT, P.; BELL, K.; BELL, I.; CHAKRABORTTY, S.; CHEN, X.; CHRAST, J.; CURADO, J.; DERRIEN, T.; DRENKOW, J.; DUMAIS, E.; DUMAIS, J.; DUTTAGUPTA, R.; FALCONNET, E.; FATUSCA, M.; FEJES-TOTH, K.; FERREIRA, P.; FOISSAC, S.; FULLWOOD, M.J.; GAO, H.; GONZALEZ, D.; GORDON, A.; GUNAWARDENA, H.; HOWALD, C.; JHA, S.; JOHNSON, R.; KAPRANOV, P.; KING, B.; KINGSWOOD, C.; LUO, O.J.; PARK, E.; PERSAUD, K.; PREALL, J.B.; RIBECA, P.; RISK, B.; ROBYR, D.; SAMMETH, M.; SCHAFFER, L.; SEE, L.H.; SHAHAB, A.; SKANCKE, J.; SUZUKI, A.M.; TAKAHASHI, H.; TILGNER, H.; TROUT, D.; WALTERS, N.; WANG, H.; WROBEL, J.; YU, Y.; RUAN, X.; HAYASHIZAKI, Y.; HARROW, J.; GERSTEIN, M.; HUBBARD, T.; REYMOND, A.; ANTONARAKIS, S.E.; HANNON, G.; GIDDINGS, M.C.; RUAN, Y.; WOLD, B.; CARNINCI, P.; GUIGÓ, R.; GINGERAS, T.R. Landscape of transcription in human cells. *Nature*, v. 489, p. 101-108, 2012.

DING, S.W.; WANG, D.W.; XIANG, Y.; XU, C.L.; XIE, H. Identification and Characterization of a Fatty Acid- and Retinoid-Binding Protein Gene (*Ar-far-1*) from the Chrysanthemum Foliar Nematode, *Aphelenchoides ritzemabosi*. *International Journal of Molecular Sciences*, v. 20, n. 22, p. 5566, 2019.

DO VALLE, C.B.; SIMIONI, C.; RESENDE, R.M.S.; JANK, L.; CHIARI, L. Melhoramento genético de *Brachiaria*. In: RESENDE, R.M.S.; DO VALLE, C.B.; JANK, L. (Ed.) *Melhoramento de forrageiras tropicais*. 1 ed. Campo Grande: Embrapa, 2008, p. 13-53.

DO VALLE, C.B.; JANK, L. E RESENDE, R.M.S. O melhoramento de forrageiras tropicais no Brasil. *Revista Ceres*, v. 56, n. 4, p. 460-472, 2009a.

DO VALLE, C.B. E PAGLIARINI, M.S. Biology, cytogenetics, and breeding of *Brachiaria*. In: SINGH, R.J. (Ed.) Genetic resources, chromosome engineering, and crop improvement. V. 5. Boca Raton, CRC Press, 2009b. p. 103-151.

DOBIN, A.; DAVIS, C.A.; SCHLESINGER, F.; DRENKOW, J.; ZALESKI, C.; JHA, S.; BATUT, P.; CHAISSON, M.; GINGERAS, T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, v. 29, n. 1, p. 15-21, 2013.

DONG, W.; LIU, J.; YU, J.; WANG, L.; ZHOU, S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One*, v. 7, n. 4, p. e35071, 2012.

DOS SANTOS, R.C.; CARNEIRO, A.C.O.; CASTRO, A.F.M.; CASTRO, R.V.O.; BIANCHE, J.J.; DE SOUZA, M.M.; CARDOSO, M.T. Correlações entre os parâmetros de qualidade da madeira e do carvão vegetal de clones de eucalipto. *Scientia Florestalis*, Piracicaba, v. 39, n. 90, p. 221-230, 2011.

DUJARDIN, M.; HANNA, W.W. Cytology and reproductive behavior of pearl millet-napiergrass hexaploides x *Pennisetum squamulatum* trispecific hybrids. *Journal of Heredity*, v. 76, n. 5, p. 382-384, 1985.

EHLTING, J.; SHIN, J.J.K.; DOUGLAS, C.J. Identification of 4-coumarate: coenzyme A ligase (4CL) substrate recognition domains. *Plant Journal*, v. 27, p. 455-465, 2001.

EMBRAPA. Anuário Leite 2018: Indicadores, tendências e oportunidades para quem vive no setor leiteiro. Juiz de Fora: Embrapa Gado de Leite, 2018.

FAN, Z.; KONG, M.; MA, L.; DUAN, S.; GAO, N.; XUQING, C.; YONGSHENG, T. Transcriptome analysis of a novel maize *bsd* C4 mutant using RNA-seq. *Plant Signaling & Behavior*, v. 15, n. 8, p. 1777374, 2020.

FARMER, A.; THIBIVILLIERS, S.; RYU, K.H.; SCHIEFELBEIN, J.; LIBAULT, M. Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in *Arabidopsis* roots at the single-cell level. *Molecular Plant*, v. 14, n. 3, p. 372-383, 2021.

FERNANDES, F.D.; CARVALHO, M.A.; RAMOS, A.K.B.; BRAGA, G.J.; FONSECA, C.E.L.; LÉDO, F.J.S.; MACHADO, J.C. Biomassa de genótipos de Capim-Elefante para produção de energia – Planaltina, DF : Embrapa Cerrados, 2020.

FERREIRA, R.P.; PEREIRA, A.V. Melhoria de forrageiras. In: BORÉM, A. (Ed.). Melhoria de espécies cultivadas. Viçosa, MG: UFV, 2005. p. 781-812.

FRASER, C.M.; CHAPPLE, C. The phenylpropanoid pathway in Arabidopsis. The Arabidopsis Book, v. 9, 2011.

GANAL, M.W.; ALTMANN, T.; RÖDER, M.S. SNP identification in crop plants. Current Opinion in Plant Biology, v. 12, p. 211-217, 2009.

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

GENG, P.; ZHANG, S.; LIU, J.; ZHAO, C.; WU, J.; CAO, Y.; FU, C.; HAN, X.; HE, H.; ZHAO, Q. MYB20, MYB42, MYB43, and MYB85 Regulate Phenylalanine and Lignin Biosynthesis during Secondary Cell Wall Formation. Plant Physiology, v. 182, p. 1272-1283, 2020.

GOERING, H.K.; VAN SOEST, P.J. Forage Fiber Analysis: Apparatus, Reagents, Procedures and Some Applications. Agricultural Handbook, Washington, D.C., p. 379, 1967.

GOMES, R.C.; FEIJÓ, G.L.D.; CHIARI, L. Evolução e Qualidade da Pecuária Brasileira. Nota técnica Embrapa Gado de Corte, Campo Grande, MS, 2017.

GOU, M.; RAN, X.; MARTIN, D.W.; LIU, C.-J. The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. Nature Plants, v. 4, p. 299-310, 2018.

GRABHERR, M.G.; HAAS, B.J.; YASSOUR, M.; LEVIN, J.Z.; THOMPSON, D.A.; AMIT, I.; ADICONIS, X.; FAN, L.; RAYCHOWDHURY, R.; ZENG, Q.; CHEN, Z.; MAUCELI, E.; HACOEN, N.; GNIRKE, A.; RHIND, N.; DI PALMA, F.; BIRREN, B.W.; NUSBAUM, C.; LINDBLAD-TOH, K.; FRIEDMAN, N.; REGEV, A. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nature Biotechnology, v. 29, n. 7, p. 644-652, 2011.

GUIMARÃES, I.C.S.B. QUANTIFICAÇÃO DE LIGNINA EM *Brachiaria spp.* MEDIDA POR LDA, LBA E NIRS. 2019. 155f. Dissertação (Mestrado em Ciências) – Universidade de São Paulo, Pirassununga.

GUO, S.G.; ZHENG, Y.; JOUNG, J.G.; LIU, S.Q.; ZHANG, Z.H.; CRASTA, O.R.; SOBRAL, B.W.; XU, Y.; HUANG, S.W.; FEI, Z.J. Transcriptome sequencing and

comparative analysis of cucumber flowers with different sex types. *BMC Genomics*, v. 11, p. 384, 2010.

GUTIÉRREZ, H.F. The American species of the genus *Cenchrus* s.l. (Poaceae, Panicoideae, Paniceae). *Darwiniana*, v. 3, n. 1, p. 114-200, 2015.

HAAS, B.J.; DELCHER, A.L.; MOUNT, S.M.; WORTMAN, J.R.; SMITH JUNIOR, R.K.; HANNICK, L.I.; MAITI, R.; RONNING, C.M.; RUSCH, D.B.; TOWN, C.D.; SALZBERG, S.L.; WHITE, O. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Research*, v. 31, n. 19, p. 5654-5666, 2003.

HABTE, E.; MUKTAR, M.S.; ABDENA, A.; HANSON, J.; SARTIE, A.M.; NEGAWO, A.T.; MACHADO, J.C.; LEDO, F.J.S.; JONES, C.S. Forage Performance and Detection of Marker Trait Associations with Potential for Napier Grass (*Cenchrus purpureus*) Improvement. *Agronomy*, v. 10, p. 542, 2020.

HACKER, J.B.; JANK, L. Breeding tropical and subtropical forage plants. In: CHERNEY, J.H. E CHERNEY, D.J.R. (Eds.) *Grass for dairy cattle*. Oxfordshire, CABI Wallingford, 1998. p. 49-71.

HANNA, W.W; SOLLENBERGER, L. E. Tropical and Subtropical Grasses. In: BARNES, R.F. (Ed.). *Forages*. 6. ed., Iowa USA: Blackwell Pub, 2007. p. 245-255.

HAO, D.C.; GE, G.B.; XIAO, P.G.; ZHANG, Y.Y.; YANG, L. The first insight into the tissue specific taxus transcriptome via Illumina second generation sequencing. *PLoS ONE*, v. 6, 2011.

HARTMANN, M.A. Plant sterols and the membrane environment. *Trends in Plant Science*, v. 3, p. 170-175, 1998.

HATFIELD, R.D.; GRABBER, J.; RALPH, J. A potential role of sinapyl p-couamate in grass lignin formation. *Plant Physiology Supplement*, v. 114, n. 86, 1997.

HERMANN, D.; FLAJOULOT, S.; JULIER, B. Sample size for diversity studies in tetraploid alfafa (*Medicago sativa*) based on codominantly coded SSR markers. *Euphytica*, v. 171, p. 441-446, 2010.

HERRERO, M.; WIRSENIUS, S.; HENDERSON, B.; RIGOLOT, C.; THORNTON, P.; HAVLIK, P.; DE BOER, I. Livestock and the Environment: What Have We Learned in the Last Decade? *Annual Review of Environment and Resources*, v. 40, p.177-202, 2015.

HIGUCHI, T. Look back over the studies of lignin biochemistry. *Journal of Wood Science*, v. 52, p. 2-8, 2006.

HIREMATH, P.J.; FARMER, A.; CANNON, S.B.; WOODWARD, J.; KUDAPA, H.; TUTEJA, R.; KUMAR, A.; BHANUPRAKASH, A.; MULAOSMANOVIC, B.; GUJARIA, N.; KRISHNAMURTHY, L.; GAUR, P.M.; KAVIKISHOR, P.B.; SHAH, T.; SRINIVASAN, R.; LOHSE, M.; XIAO, Y.L.; TOWN, C.D.; COOK, D.R.; MAY, G.D.; VARSHNEY, R.K. Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnology Journal*, v. 9, p. 922-931, 2011.

HIMMEL, M.E.; DING, S.; JOHNSON, D.K.; ADNEY, W.S.; NIMLOS, M.R.; BRADY, H.W.; FOUST, T.D. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science*, v. 315, p. 804-807, 2007.

HSU, T.W.; SHIH, H.C.; KUO, C.C.; CHIANG, T.Y.; CHIANG, Y.C. Characterization of 42 microsatellite markers from poison ivy, *Toxicodendron radicans* (Anacardiaceae). *International Journal of Molecular Sciences*, v. 14, p. 20414-20426, 2013.

HUANG, W.; GUO, Y.; DU, W.; ZHANG, X.; LI, A.; MIAO, X. Global transcriptome analysis identifies differentially expressed genes related to lipid metabolism in Wagyu and Holstein cattle. *Scientific Reports*, v. 7, n. 1, p. 5278, 2017.

HUBER, G.W.; IBORRA, S.; CORMA, A. Synthesis of transportation fuels from biomass: chemistry, catalysis and engineering. *Chemical reviews*, v. 106, n. 9, p. 4044-4098, 2006.

HWANG, T-Y.; GWAK, B.S.; SUNG, J.; KIM, H-S. Genetic Diversity Patterns and Discrimination of 172 Korean Soybean (*Glycine max* (L.) Merrill) Varieties Based on SSR Analysis. *Agriculture*, v. 10, 2020.

IPEK, M.; IPEK, A.; SIMON, P.W. Comparison of AFLPs, RAPD markers, and isozymes for diversity assessment of garlic and detection of putative duplicates in germplasm collections. *Journal of the American Society for Horticultural Science*, v. 128, p. 246-252, 2003.

IQBAL, M.Z.; JAMIL, S.; MEHMOOD, A.; SHAHZAD, R. Identification of Seven olive varieties using RAPD molecular markers. *Journal of Agricultural Research*, v. 57, p. 7-14, 2019.

IQBAL, M.Z.; JAMIL, S.; SHAHZAD, R.; BILAL, K.; QAISAR, R.; NISAR, A.; KANWAL, S.; BHATTI, M.K. DNA Fingerprint of crops and its application in the field of plant breeding. *Journal of Agricultural Research*, v. 59, n. 1, p. 13-28, 2021.

JAHANGEER, M.; FATIMA, R.; ASHIQ, M.; BASHARAT, A.; QAMAR, S.A.; BILAL, M.; IQBAL, H.M.N. Therapeutical and Biomedical Potentialities of Terpenoids – A Review. *Journal of Pure and Applied Microbiology*, 2021.

JAMIL, S.; SHAHZAD, R.; YASMEEN, E.; RAHMAN, S.U.; YOUNAS, M.; IQBAL, M.Z. DNA fingerprint of Pakistani maize hybrids and parental lines using simple sequence repeat markers. *Pakistan Journal of Botany*, v. 52, 2020a.

JAMIL, S.; SHAHZAD, R.; KANWAL, S.; YASMEEN, E.; RAHMAN, S.U.; IQBAL, M.Z. DNA Fingerprint and Population Structure of Sate Palm Varieties Grown in Punjab Pakistan using Simple Sequence Repeat Markers. *International Journal of Agriculture and Biology*, v. 23, p. 943-950, 2020b.

JANK, L.; CARVALHO, P.F. E DO VALLE, C.B. New grasses and legumes: advances and perspectives for the tropical zones of Latin America. In: REYNOLDS, S.G. E FRAME, J. (Org.) *Grasslands: developments, opportunities, perspectives*. Roma, FAO, Índia, Science Publisher, 2005. p. 55-79.

JARADAT, A.A. Genetic resources of energy crops: biological systems to combat climate change. *Australian Journal Crop Science*, v. 4, n. 5, p. 309-323, 2010.

JIANG, L.F.; ZHANG, X.Q.; MA, X.; HUANG, L.K.; XIE, W.G.; MA, Y.M.; ZHAO, Y.F. Identification of orchardgrass (*Dactylis glomerata* L.) cultivars by using simple sequence repeat markers. *Genetics and Molecular Research*, v. 12, p. 5111-5123, 2013.

JORDON-THADEN, I.E.; CHANDERBALI, A.S.; GITZENDANNER, M.A.; SOLTIS, D.E. Modified CTAB and TRizol protocols improve RNA extraction from chemically complex Embryophyta. *Applications in Plant Sciences*, v. 3, n. 5, 2015.

JOZWIAK, A.; SONAWANE, P.D.; PANDA, S.; GARAGOUNIS, C.; PAPADOPOULOU, K.K.; ABEBIE, B.; MASSALHA, H.; ALMEKIAS-SIEGL, E.; SCHERF, T.; AHARONI, A. Plant terpenoid metabolism co-opts a component of the cell wall biosynthesis machinery. *Nature Chemical Biology*, v. 16, p. 740-748, 2020.

JUNGMANN, L.; VIGNA, B.B.Z.; PAIVA, J.; SOUSA, A.C.B.; DO VALLE, C.B.; LABORDA, P.R.; ZUCCHI, M.I.; DE SOUZA, A.P. Development of microsatellite markers for *Brachiaria humidicola* (Rendle) Schweick. *Conservation Genetic Resources*, 2009.

JUNGMANN, L.; VIGNA, B.B.Z.; BOLDRINI, K.R.; SOUSA, A.C.B.; DO VALLE, C.B.; RESENDE, R.M.S.; PAGLIARINI, M.S.; ZUCCHI, M.I.; DE SOUZA, A.P. Genetic diversity and population structure analysis of the tropical pasture grass *Brachiaria humidicola* based on microsatellites, cytogenetics, morphological traits, and geographical origin. NCR Research Press, 2010.

KANDEL, R.; SINGH, H.P.; SINGH, B.P.; HARRIS-SHULTZ, K.; ANDERSON, W.F. Assessment of Genetic Diversity in Napier Grass (*Pennisetum purpureum* Schum.) using Microsatellite, Single-Nucleotide Polymorphism and Insertion-Deletion Markers from Pearl Millet (*Pennisetum glaucum* [L.] R. Br.). Plant Molecular Biology Reporter, v. 34, p. 265-272, 2016.

KANEHISA, M.; GOTO, S.; SATO, Y.; FURUMICHI, M.; TANABEL, M. KEGG for integration and interpretation of large-scale molecular datasets. Nucleic Acids Research, v. 40, p. 109-114, 2012.

KARIA, C.T.; DUARTE, J.B.; DE ARAÚJO, A.C.G. Desenvolvimento de Cultivares do gênero *Brachiaria* (trin.) Griseb. No Brasil. Documentos 163 – Planaltina, DF: Embrapa Cerrados, 2006.

KAUR, S.; COGAN, N.O.I.; PEMBLETON, L.W.; SHINOZUKA, M.; SAVIN, K.W.; MATERNE, M.; FOSTER, J.W. Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. BMC Genomics, v. 12, p. 265, 2011.

KELLER-GREIN, G.; MAAS, B.L.; HANSON, J. Natural variation in *Brachiaria* and existing germplasm collections. In: MILES, J.W.; MAAS, B.L.; VALLE, C.B. do (Ed.) *Brachiaria: biology, agronomy, and improvement*. Cali: CIAT, 1996. cap. 2, p. 16-42.

KERING, M.K.; GURETZKY, J.; FUNDERBURG, E.; MOSALI, J. Effect of Nitrogen Fertilizer Rate and Harvest Season on Forage Yield, Quality, and Macronutrient Concentrations in Midland Bermuda Grass, Commun. Soil Science and Plant Analysis, v. 42, p. 1958-1971, 2011.

KO, J.K.; LEE, J.H.; JUNG, J.H.; LEE, S.M. Recent advances and future directions in plant and yeast engineering to improve lignocellulosic biofuel production. Renewable and Sustainable Energy Reviews, v. 134, p. 110390, 2020.

KONG, W.; ZHANG, C.; QIANG, Y.; ZHONG, H.; ZHAO, G.; LI, Y. Integrated RNA-Seq analysis and Meta-QTLs mapping provide insights into cold stress response in rice seedlings roots. *International Journal of Molecular Sciences*, v. 21, n. 13, p. 4615, 2020.

KOUKOL, J.; CONN, E. E. The metabolism of aromatic compounds in higher plants IV. Purification and properties of the phenylalanine deaminase of *Hordeum vulgare*. *Journal of Biological Chemistry*, v. 236, n. 10, p. 2692-2698, 1961.

KROGH, A.; LARSSON, B.; VON HEIJNE, G.; SONNHAMMER, E.L. Predicting transmembrane protein topology with a hidden Markov level: application to complete genomes. *Journal of Molecular Biology*, v. 305, n. 3, p. 567-580, 2001.

KUWI, S.O.; KYALO, M.; MUTAI, C.K.; MWILAWA, A.; HANSON, J.; DJIKENG, A.; GHIMIRE, S.R. Genetic diversity and population structure of *Urochloa* grass accessions from Tanzania using simple sequence repeat (SSR) markers. *Brazilian Journal of Botany*, v. 41, p. 699-709, 2018.

LALAK-KANCZUGOWSKA, J.; MARTYNIAK, D.; KASPRZYCKA, A.; ŻUREK, G.; MORON, W.S.; CHMIELEWSKA, M.; WIACEK, D.; TY, J. Comparison of selected parameters of biomass and coal. *International Agrophysics*, v. 30, n. 4, p. 475-482, 2016.

LANDER, E.S.; BOTSTEIN, S. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, v. 121, p. 185, 1989.

LANDJEVA, S.; KORZUN, V.; BÖRNER, A. Molecular markers: actual and potential contributions to wheat genome characterization and breeding. *Euphytica*, v. 156, p. 271-296, 2007.

LAVHALE, S.G.; KALUNKE, R.M.; GIRI, A.P. Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants. *Planta*, v. 248, p. 1063-1078, 2018.

LEE, C.H.; TENG, Q.; HUANG, W.L.; ZHONG, R.Q.; YE, R.Q.; YE, Z.H. Down-regulation of PoGT47C expression in poplar results in a reduced glucuronoxylan content and an increased wood digestibility by cellulase. *Plant & Cell Physiology*, v. 50, n. 6, p. 1075-1089, 2009.

LEE, M.A.; DAVIS, A.P.; CHAGUNDA, M.G.G.; MANNING, P. Forage quality declines with rising temperatures, with implications for livestock production and methane emissions. *Biogeosciences*, v. 14, p. 1403-1417, 2017.

LEWANDOWSKI, I.; SCURLOCK, J.M.O.; LINDVALL, E.; CHRISTOU, M. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass Bioenergy*, v. 25, p. 335-361, 2003.

LI, M.; PU, Y. RAGAUSKAS, A.J. Current understanding of the correlation of lignin structure with biomass recalcitrance. *Frontiers in Chemistry*, v. 4, p. 45, 2016a.

LI, R.; CHEN, J. Planning the next-generation biofuel crops based on soil-water constraints. *Biomass and Bioenergy*, v. 115, p. 19-26, 2018.

LI, W.; NOTANI, D.; ROSENFELD, M.G. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nature Reviews Genetics*, v. 17, p. 207-223, 2016b.

LI, Y.; HU, L.J.; WU, X.Q.; YE, J.R. A *Bursaphelenchus xylophilus* effector, Bx-FAR-1, suppresses plant defense and affects nematode infection of pine trees. *European Journal of Plant Pathology*, v. 157, p. 637-650, 2020.

LIU, Q.; LUO, L.; ZHENG, L. Lignins: biosynthesis and biological functions in plants. *International Journal in Molecular Sciences*, v. 19, p. 335, 2018.

LIU, Q.; LU, Z.; HE, W.; LI, F.; CHEN, W.; LI, C.; CHAO, Z.; TIAN, E. Development and characterization of 16 novel microsatellite markers by Transcriptome sequencing for *Angelica dahurica* and test for cross-species amplification. *BMC Plant Biology*, v. 20, p. 152, 2020a.

LIU, X.H.; LYU, Y.S.; YANG, W.; YANG, Z.T.; LU, S.J.; LIU, J.X. A membrane-associated NAC transcription factor OsNTL3 is involved in thermotolerance in rice. *Plant Biotechnology Journal*, n. 18, n. 5, p. 1317-1329, 2020b.

LOVE, M.I.; HUBER, W.; ANDERS, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, v. 15, p. 550, 2014.

LOVETT, A.A.; SÜNNENBERG, G.M.; RITCHER, G.M.; DAILEY, A.G.; RICHE, A.B.; KARP, A. Land use implications of increased biomass production identified by gis-based suitability and yield mapping for miscanthus in England. *BioEnergy Research*, v. 2, p. 17-28, 2009.

LU, J.; ZHANG, Y.; DIAO, X.; YU, K.; DAI, X.; QU, P.; CRABBE, J.C.; ZHANG, T.; QIAO, Q. Evaluation of genetic diversity and population structure of *Fragaria nilgerrensis* using EST-SSR markers. *Gene*, v. 796, p. 145791, 2021.

LYONS, R.K.; MACHEN, R.; FORBES, T.D.A. Why range forage quality changes. Bulletin/Texas Agricultural Extension Service; no. 6036., 1996.

MACKILL, D.J.; NGUYEN, H.T.; ZHANG, J. Use of molecular markers in plant improvement programs for rainfed lowland rice. *Field Crops Research*, v. 64, p. 177-185, 1999.

MADRITSCH, S.; BURG, A.; SEHR, E.M. Comparing de novo transcriptome assembly tools in di- and autotetraploid non-model plant species. *BMC Bioinformatics*, v. 22, p. 146, 2021.

MANARA, N.T.F. Citogenética de variedades de capim elefante (*Pennisetum purpureum* Schum.). 1973. 63 p. Dissertação (Mestrado em Genética) – Escola Superior de Agricultura Luiz Queiroz, Piracicaba.

MAPA (Ministério da Agricultura, Pecuária e Abastecimento). Projeções do Agronegócio – Brasil 2020/21 a 2030/31. Brasília, 2021.

MARGUERAT, S.; BÄHLER, J. RNA-Seq: from technology to biology. *Cellular and Molecular Life Sciences*, v. 67, p. 569-579, 2010.

MARSCHNER, M. Mineral nutrition of higher plants, 2 ed, Academic Press, London, 1995.

MATTHEW, C.; VAN LOO, E.N.; THOM, E.R.; DAWSON, L.A.; CARE, D.A. Understanding shoot and root development. In: GOMIDE, J.A.; MATTOS, W.R.S.; DA SILVA, S.C.; FILHO, A.M.C. (Eds.) Proceedings of the XIX International Grassland Congress. São Paulo, Brasil, 2001, p. 19-27.

MCKENDRY, P. Energy production from biomass (Part 1): overview of biomass. *Bioresource Technology*, v. 83, n. 1, p. 37-46, 2002.

MILES, J.W.; DO VALLE, C.B. Manipulation of apomixis in *Brachiaria* breeding. In: MILES, J.W.; MAASS, B.L. E DO VALLE, C.B. (Eds.) *Brachiaria: biology, agronomy, and improvement*. CIAT/Brasília:EMBRAPA-CNPQC, 1996. p. 164-177.

MILES, J.W. Apomixis for cultivar development in tropical forage grasses. *Crop Science*, v. 47, p. 238-249, 2007.

MILES, J.W.; MAASS, B.L.; DO VALLE, C.B.; KUMBLE, V. *Brachiaria: biology, agronomy and improvement*. Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Campo Grande, p. 288, 1996.

MILES, J.W.; DO VALLE, C.B.; RAO, I.M.; EUCLIDES, V.P. *Brachiaria* grasses. In: MOSER, L.E.; BURSON, B.L.; SOLLENBERGER, L.E. Warm-season (C4) grasses. American Society of Agronomy, Madison, Wisconsin, 2004, p. 745-783.

MONRROY, M.; GUTIÉRREZ, D.; MIRANDA, M.; HERNÁNDEZ, K.; GARCÍA, J.R. Determination of *Brachiaria* spp. forage quality by near-infrared spectroscopy and partial least squares regression. Journal of the Chilean Chemical Society, v. 62, n. 2, p. 3472-3477, 2017.

MORALES-QUINTANA, L.; BELTRÁN, D.; MENDEZ-YAÑEZ, Á.; VALENZUELA-RIFFO, F.; HERRERA, R.; MOYA-LEÓN, M.A. Characterization of FcXTH2, a Novel Xyloglucan Endotransglycosylase/Hydrolase Enzyme of Chilean Strawberry with Hydrolase Activity. International Journal of Molecular Sciences, v. 21, 2020.

MORGAN, J.L.W.; MCNAMARA, J.T.; FISCHER, M.; RICH, J.; CHEN, H.M.; WITHERS, S.G.; ZIMMER, J. Observing cellulose biosynthesis and membrane translocation in *crystallo*. Nature, v. 531, p. 329-334, 2016.

MORGANTE, M.; HANAFEY, M.; POWELL, W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetics, v. 30, n. 2, p. 1-16, 2002.

MORIN, R.D.; BAINBRIDGE, M.; FEJES, A.; HIRST, M.; KRZYWINSKI, M.; PUGH, T.J.; McDONALD, H.; VARHOL, R.; JONES, S.J.M.; MARRA, M.A. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques, v. 45, p. 81-94, 2008.

MORRIS, K.V.; MATTICK, J.S. The rise of regulatory RNA. Nature Reviews Genetics, v. 15, p. 423-437, 2014.

MOROZOVA, O.; MARRA, M. Applications of next-generation sequencing technologies in functional genomics. Genomics, v. 92, p. 255-264, 2008.

NABAVI, S.M.; ŠAMEC, D.; TOMCZYK, M.; MILELLA, L.; RUSSO, D.; HABTEMARIAN, S.; SUNTAR, I.; RASTRELLI, L.; DAGLIA, M.; XIAO, J.; GIAMPIERI, F.; BATTINO, M.; SOBARZO-SANCHEZ, E.; NABAVI, S.F.; YOUSEFI, B.; JEANDET, P.; XU, S.; SHIROOIE, S. Flavonoid biosynthetic pathways in plants: Versatile targets for metabolic engineering. Biotechnology Advances, v. 38, 2018.

NADEEM, M.A.; NAWAZ, M.A.; SHAHID, M.Q.; DOĞAN, Y.; COMENTPAY, G.; YILDIZ, M.; HATIPOĞLU, R.; AHMAD, F.; ALSALEH, A.; LABHANE, N.; ÖZKAN, H.; CHUNG, G.; BALOCH, F.S. 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.

NAGEGOWDA, D.A.; GUPTA, P. Advances in biosynthesis, regulation, and metabolic engineering of plant specialized terpenoids. *Plant Science*, v. 294, p. 110457, 2020.

NAKAMO, Y.; YAMAGUCHI, M.; ENDO, H.; REJAB, N.A.; OHTANI, M. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Frontiers in Plant Science*, v. 6, p. 288, 2015.

NASCIMENTO JÚNIOR, D. Informações sobre plantas forrageiras. Viçosa, MG: UFV-Imprensa Universitária, 1981. 56p.

NCBI, 2021. Disponível em: <http://www.ncbi.nlm.nih.gov/gquery/?term=poaceae>. Acesso em 03 agosto 2021.

NEWMAN, Y.C.; LAMBERT, B.; MUIR, J.P. Defining Forage Quality. Texas Cooperative Extension, 2007.

NIE, G.; HUANG, T.; MA, X.; HUANG, L.; PENG, Y.; YAN, Y.; LI, Z.; WANG, X.; ZHANG, X. Genetic variability evaluation and cultivar identification of tetraploid annual ryegrass using SSR markers. *PeerJ*, v. 7, 2019.

NOVAES, E.; DROST, D.R.; FARMERIE, W.G.; PAPPAS, G.J.; GRATTAPAGLIA, D.; SEDEROFF, R.R.; KIRST, M. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics*, v. 9, p. 312, 2008.

NUÑEZ, J.; AREVALO, A.; KARWAT, H.; EGENOLF, K.; MILES, J.; CHIRINDA, N.; CADISH, G.; RASCHE, F.; RAO, I.; SUBBARAO, G.; ARANGO, J. Biological nitrification inhibition activity in a soi-grown biparental population of the forage grass, *Brachiaria humidicola*. *Plant Soil*, v. 426, p. 401-411, 2018.

O'HANLON, P.C.; PEAKALL, R.; BRIESE, D.T. A review of new PCR-based genetic markers and their utility to weed ecology. *Weed Research*, v. 40, p. 239-254, 2000.

OHTANI, M.; DEMURA, T. The quest for transcriptional hubs if lignin biosynthesis: beyond the NAC-MYB-gene regulatory network model. *Current Opinion in Biotechnology*, v. 56, p. 82-87, 2019.

OLIVEIRA, F.A.; VIGNA, B.B.Z.; SILVA, C.C.; FÁVERO, A.P.; MARTA, F.P.; AZEVEDO, A.L.S.; DE SOUZA, A.P. Apomixis-related genes identified in a coexpression network of the neotropical grass *Paspalum notatum*. bioRxiv, p. 369280, 2018.

OLIVEIRA, D.M.; MOTA, T.R.; GRANDIS, A.; DE MORAIS, G.R.; DE LUCAS, R.C.; POLIZELI, M.L.T.L.; MARCHIOSI, R.; BUCKRIDGE, M.S.; FERRARESE-FILHO, O.; DOS SANTOS, W.D. Lignin plays a role in determining biomass recalcitrance in forage grasses. *Renewable Energy*, v. 147, p. 2206-2217, 2020.

ÖNNERUD, H.; ZHANG, L.; GELLERSTEDT, G. HENRIKSSON, G. Polymerization of monolignols by redox shuttle-mediated enzymatic oxidation: a new model in lignin biosynthesis. *Plant Cell*, v. 14, p. 1953-1962, 2002.

ORTIZ, R. Molecular plant breeding. *Crop Science*, v. 50, p. 2196-2197, 2010.

OUTLOOK BRASIL 2022. Disponível em: <http://icna.org.br/sites/default/files/artigo/Outlook%20Brasil%202022.pdf>. Acesso em 28 maio 2019.

PAN, L.P.; HUANG, T.; YANG, Z.; TANG, L.; CHENG, Y.; WANG, J.; MA, X.; ZHANG, X. EST-SSR marker characterization based on RNA-sequencing of *Lolium multiflorum* and cross transferability to related species. *Molecular Breeding*, v. 38, p. 80, 2018.

, Y.S.; PARK, J.S.; LEE, S.; JUNG, S.H.; KIM, S.K.; RYU, C.M. Simultaneous profiling of *Arabidopsis thaliana* and *Vibrio vulnificus* MO6-24/O transcriptomes by dual RNA-seq analysis. *Computational and Structural Biotechnology Journal*, v. 19, p. 2084-2096, 2021.

PATRO, R.; DUGGAL, G.; LOVE, M.I.; IRIZARRI, R.A.; KINGSFORD, C. Salmon provides fast and bias-aware quantification of transcription expression. *Nature Methods*, v. 14, p. 417-419, 2017.

PEREIRA, A.V.; DO VALLE, C.B. E MILES, J.W. Melhoramento de forrageiras tropicais. In: NASS, L.L.; VALOIS, A.C.C.; DE MELO, I.S.; VALADARES-INGLIS, M.C. (Eds.) Recursos energéticos & melhoramento – plantas. Rondonópolis, Fundação MT. p. 549-602, 2001.

PEREIRA, A.V.; SOUZA SOBRINHO, F.; SOUZA, F.H.D.; LÉDO, F.J.S. Tendências do melhoramento genético e produção de sementes forrageira no Brasil. In: SIMPÓSIO SOBRE ATUALIZAÇÃO EM GENÉTICA E MELHORAMENTO DE PLANTAS, v. 4, 2003, Lavras. Anuais... Lavras: UFLA, 2003. p. 36-63.

PEREIRA, A.V.; AUAD, A.M.; LÉDO, F.J.S.; BARBOSA, S. *Pennisetum purpureum*. In: FONSECA, D.M.; MARTUSCELLO, J.A. (Ed.) Plantas forrageiras. Viçosa, MG: Ed. UFV, 2010. cap. 6, p. 197-219.

PEREIRA, A.V.; MORENZ, M.J.F.; LÉDO, F.J.S.; FERREIRA, R.P. Capim-elefante: Versatilidades de usos na pecuária de leite. In: VILELA, D.; FERREIRA, R.P.; FERNANDES, E.N.; JUNTOLLI, F.V. (Org.) Pecuária de leite no Brasil. 1. ed. Brasília, DF: Embrapa, 2016a. Parte 2, cap. 3, p. 187-211.

PERTEA, M.; PERTEA, G.M.; ANTONESCU, C.M.; CHANG, T.C.; MENDELL, J.T.; SALZBERG, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 2015.

PESSOA-FILHO, M.; MARTINS, A.M.; FERREIRA, M.E. Molecular dating of phylogenetic divergence between *Urochloa* species based on complete chloroplast genomes. *BMC Genomics*, v. 18, 2017.

PESTANA-CALSA, M.C.; PACHECO, C.M.; DE CASTRO, R.C.; DE ALMEIDA, R.R.; DE LIRA, N.P.V.; CALSA JUNIOR, T. Cell wall, lignin and fatty acid-related transcriptome in soybean: Achieving gene expression patterns for bioenergy legume. *Genetics and Molecular Biology*, v. 32, n.1, p. 322-330, 2012.

PHILIPPE, F.; PELLOUX, J.; RAYON, C. Plant pectin acetylesterase structure and function: new insights from bioinformatic analysis. *BMC Genomics*, v. 18, n. 1, p. 1-18, 2017.

POPP, J.; HARANGI-RAKOS, M.; GABNAI, Z.; ANTAL, G.; BAI, A. Biofuels and their co-products as livestock feed: global economic and environmental implications. *Molecules*, v. 21, p. 285, 2016.

PRONOBIS, M. The influence of biomass co-combustion on boiler fouling and efficiency. *Fuel*, v. 85, n. 4, p. 474-480, 2006.

PROTÁSIO, T.P.; BUFALINO, L.; TONOLI, G.H.D.; COUTO, A.M.; TRUGILHO, P.F.; GUIMARÃES JÚNIO, M. Relação entre o poder calorífero superior e os componentes elementares e minerais da biomassa vegetal. *Pesquisa Florestal Brasileira*, v. 31, n. 66, p. 113-122, 2011.

PROVAN, J.; POWELL, W.; HOLLINGSWORTH, P.M. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution*, v. 16, p. 142-147, 2001.

PUNTA, M.; COGGILL, P.C.; EBERHARDT, R.Y.; MISTRY, J.; TATE, J.; BOURSNELL, C.; PANG, K.; FORSLUND, K.; CERIC, G.; CLEMENTS, J.; HEGER, A.; HOLM, L.; SONNHAMMER, E.L.L.; EDDY, S.R.; BATEMAN, A.; FINN, R.D. The Pfam protein families database. *Nucleic Acids Research*, v. 40, n. D1, p. D290-D301, 2012.

PURTY, R.S.; SACHAR, M.; CHATTERJEE, S. Structural and Expression Analysis of Salinity Stress Responsive Phosphoserine Phosphatase from *Brassica juncea* (L.). *Journal of Proteomics & Bioinformatics*, v. 10, n. 4, 2017.

PURUSHOTHAN, P.; HO, R.; ZIMMER, J. Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science*, v. 369, p. 1089-1094, 2020.

QIU, F.; BACHLE, S.; ESTES, R.; DUVALL, M.R.; NIPPERT, J.B.; UNGERER, M.C. Transcriptional responses to water stress and recovery in a drought-tolerance fescue wild grass (*Festuca ovina*; Poaceae). *Genome*, v. 64, n. 1, p. 15-27, 2020.

QUENTIN, M.; ALLASIA, V.; PEGARD, A.; ALLAIS, F.; DUCROT, P.H.; FAVERY, B.; LEVIS, C.; MARTINET, S.; MASUR, C.; PONCHET, M.; ROBY, D.; SCHLAICH, N.L.; JOUANIN, L.; KELLER, H. Imbalanced lignin biosynthesis promotes the sexual reproduction of homothallic oomycete pathogens. *PLoS Pathogens*, v. 5, n. 1, p. e1000264, 2009.

RAES, J.; ROHDE, A.; CHRISTENSEN, J.H.; VAN DE PEER, Y.; BOERJAN, W. Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiology*, v. 133, p. 1051-1071, 2003.

RAFFRENATO, E.; FIEVISOHN, R.; COTANCH, K.W.; GRANT, R.J.; CHASE, L.E.; VAN AMBURGH, M.E. Effect of lignin linkages with other plant cell wall components on in vitro and in vivo neutral detergent fiber digestibility and rate of digestion of grass forages. *Journal of Dairy Science*, v. 100, p. 8119-8131, 2016.

RAJENDRAKUMAR, P.; BISWAL, A.K.; BALACHANDRAN, S.M.; SRINIVASARAO, K.; SUNDARAM, R.M. Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. *Bioinformatics*, v. 23, n. 1, p. 1-4, 2007.

RALPH, J.; LAPIERRE, C.; BOERJAN, W. Lignin structure and its engineering. *Current Opinion in Biotechnology*, v. 56, p. 240, 249, 2019.

RAO, I.M.; KERRIDGE, P.C. E MACEDO, M. Nutritional requirements of *Brachiaria* and adaptation to acid soils. In: MILES, J.W.; MAASS, B.L.; DO VALLE, C.B. (Eds.) *Brachiaria: The biology, agronomy, and improvement*. CIAT, Cali, 1996, p. 53-71.

RAO, X.; CHEN, X.; SHEN, H.; MA, Q.; LI, G.; TANG, Y.; PENA, M.; YORK, W.; FRAZIER, T.P.; LENAGHAN, S.; XIAO, X.; CHEN, F.; DIXON, R.A. Gene regulatory networks for lignin biosynthesis in switchgrass (*Panicum virgatum*). *Plant Biotechnology Journal*, v. 17, n. 3, p. 580-593, 2019.

REIS, G.B. Mecanismos da eliminação cromossômica em híbridos de *Pennisetum purpureum* e *P. glaucum* e relações genômicas em *Pennisetum* spp. 2016. 96p. Tese Acadêmica (Doutorado em Genética e Melhoramento de Plantas) - Universidade Federal de Lavras, Lavras.

REIS, R.A.; BARBERO, R.P.; HOFFMANN, A. Impactos da qualidade da forragem em sistemas de produção de bovinos de corte. *Informe Agropecuário*, Belo Horizonte, v. 37, n. 292, p. 36-53, 2016.

RISSE, D.; SCHWARTZ, K.; SHERLOCK, G.; DUDOIT, S. GC-content normalization for RNA-seq data. *BMC Bioinformatics*, v. 12, p. 480, 2011.

RISSE, D.; NGAI, J.; SPEED, T.P.; DUDOIT, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology*, v. 32, p. 896-902, 2014.

ROCHA, J.R.A.S.C.; MARÇAL, T.S.; SALVADOR, F.V.; SILVA, A.C.; CARNEIRO, P.C.S.; RESENDE, M.D.V.; CARNEIRO, J.C.; AZEVEDO, A.L.S.; PEREIRA, J.F.; MACHADO, J.C. Unraveling candidate genes underlying biomass digestibility in elephant grass (*Cenchrus purpureus*). *BMC Plant Biology*, v. 19, p. 548, 2019.

RODRIGUEZ, F.; PEREZ, T.; HAMMER, S.E.; ALBORNOZ, J.; DOMINGUEZ, A. Integrating phylogeographic patterns of microsatellite and mtDNA divergence to infer the evolutionary history of chamois (genus *Rupicapra*). *BMC Ecology and Evolution*, v. 10, p. 222, 2010.

ROGOWSKA, A.; SKAZIEL, A. The role of sterols in plant response to abiotic stress. *Phytochemistry Reviews*, v. 19, p. 1525-1538, 2020.

ROSLER, J.; KREKEL, F.; AMRHEIN, N.; SCMID, J. Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiology*, v. 113, p. 175-179, 1997.

ROYO, J.B.; CABELLO, F.; MIRANDA, S.; GOGORCENA, Y.; GONZALEZ, J.; MORENO, S.; ITOIZ, R.; ORTIZ, J.M. The use of isoenzymes in characterization of grapevines (*Vitis vinifera* L.). Influence of the environment and time of sampling. *Scientia Horticulturae*, v. 69, p. 145-155, 1997.

RUEDA, J.A.; ORTEGA-JIMÉNEZ, E.; HERNÁNDEZ-GARAY, A.; ENRÍQUEZ-QUIROZ, J.; GUERRERO-RODRÍGUEZ, J.; QUERO-CARRILLO, A.R. Growth, yield, fiber content and lodging resistance in eight varieties of *Cenchrus purpureus* (Schumach.) Morrone intended as energy crop. *Biomass and Bioenergy*, v. 88, p. 50-65, 2016.

RUEDA, J.A. GUERRERO-RODRÍGUEZ, J.; RAMÍREZ-ORDOÑES, S.; AGUILAR-MARTÍNEZ, C.U.; HERNÁNDEZ-MONTIEL, W.; ORTEGA-JIMÉNEZ, E. Morphological composition and fiber partitioning along regrowth in elephant grass CT115 intended for ethanol production. *Scientific Reports*, v. 10, p. 15118, 2020.

SAHRAEIAN, S.M.E.; MOHIYUDDIN, M.; SEBRA, R.; TILGNER, H.; AFSHAR, P.T.; AU, K.F.; ASADI, N.B.; GERSTEIN, M.B.; WONG, W.W.; SNYDER, M.P.; SCHADT, E.; LAM, H.Y.K. Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nature Communications*, v. 8, n. 59, 2017.

SALEM, H.H.; ALI, B.A.; HUANG, T.H.; QIN, D.N.; WANG, X.M.; XIE, Q.D. Use of random amplified polymorphic DNA analysis for economically important food crops. *Journal of Intergrative Plant Biology*, v. 49, p. 1670-1680, 2007.

SALGADO, L.R.; LIMA, R.; SANTOS, B.F.; SHIRAKAWA, K.T.; VILELA, M.A.; ALMEIDA, N.F.; PEREIRA, R.M.; NEPOMUCENO, A.L.; CHIARI, L. De novo RNA sequencing and analysis of the transcriptome of signalgrass (*Urochloa decumbens*) roots exposed to aluminum. *Plant Growth Regulation*, v. 83, p. 157-170, 2017.

SAMPEDRO, J.; COSGROVE, D.J. The expansin superfamily. *Genome Biology*, v. 6, p. 242, 2005.

SATTURU, V.; RANI, D.; GATTU, S.; MD, J.; MULINTI, S.; NAGIREDDY, R.K.; ERUVURI, R.; YANDA, R. DNA Fingerprint for Identification of Rice Varieties and Seed Genetic Purity Assessment. *Agricultural Research*, v. 7, n. 4, p. 379-390, 2018.

SCHLIESKY, S.; GOWIK, U.; WEBER, A.P.M.; BRÄUTIGAM, A. RNA-seq assembly – are we there yet? *Frontiers in Plant Science*, v. 3, 2012.

SCHUTYSER, W.; RENDERS, T.; VAN DER BOSCH, S.; KOELEWIJN, S.-F.; BECKHAM, G.T.; SELS, B.F. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation and upgrading. *Royal Society of Chemistry*, v. 4, p. 852, 2018.

SEWALT, V.J.H.; NI, W.T.; JUNG, H.G.; DIXON, R.A. Lignin impact on fiber degradation: increased enzymatic digestibility of genetically engineered tobacco (*Nicotiana tabacum*)

stems reduced in lignin content. *Journal of Agricultural and Food Chemistry*, v. 45, p. 1977-1983, 1997.

SIBOUT, R.; EUDES, A.; POLLET, B.; GOUJON, T.; MILA, I.; GRANIER, F.; SEGUIN, A.; LAPIERRE, C.; JOUANIN, L. Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenases in *Arabidopsis*. Isolation and characterization of the corresponding mutants. *Plant Physiology*, v. 132, p. 848-860, 2003.

SINGH, R.B.; SINGH, B.; SINGH, R.K. Identification of elite Indian sugarcane varieties through DNA fingerprint using genic microsatellite markers. *Vegetos*, v. 32, p. 547-555, 2019.

SOUZA SOBRINHO, F.; PEREIRA, A.V.; LEDO, F.J.S.; BOTREL, M.A.; OLIVEIRA, J.S.; XAVIER, D.F. Avaliação agronômica de híbridos interespecíficos entre capim-elefante e milho. *Pesquisa Agropecuária Brasileira*, v. 40, n. 9, p. 873-880, 2005.

SOUZA, J.S.; CHIARI, L.; SIMEÃO, R.M.; VILELA, M.M.; SALGADO, L.R. Development, Validation and Characterization of Genic Microsatellite Markers in *Urochloa* Species. *American Journal of Plant Sciences*, v. 9, p. 281, 295, 2018.

STAPT, O.; HUBBARD, C.E. *Pennisetum*. In: PRAIN, D. (Ed.). *Flora of tropical Africa*. Ashford: Reeve, 1934. v. 9, p. 954-1070.

STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. *Nature Reviews Genetics*, v. 20, p. 631-656, 2019.

STEWART, J.J.; AKIYAMA, T.; CHAPPLE, C.; RALPH, J.; MANSFIELD, S.D. The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiology*, v. 150, p. 621-635, 2009.

SU, Y.; ZHANG, P.; SU, Y. An overview of biofuels policies and industrialization in the major biofuel producing countries. *Renewable and Sustainable Energy Reviews*, v. 50, p. 991-1003, 2015.

SU, Q.; ZHANG, X.; ZHANG, W.; ZHANG, N.; SONG, L.; LIU, L.; XUE, X.; LIU, G.; LIU, J.; MENG, D.; ZHI, L.; JI, J.; ZHAO, X.; YANG, C.; TONG, Y.; LIU, Z.; LI, J. QTL Detection for Kernel Size and Weight in Bread Wheat (*Triticum aestivum* L.) Using a High-Density SNP and SSR-Based Linkage Map. *Frontiers in Plant Science*, v. 9, p. 1484, 2018.

SUMARANI, G.; PILLAI, S.V.; HARISANKAR, P.; SUNDARESAN, S. Isozyme analysis of indigenous cassava germplasm for identification of duplicates. *Genetic Resources and Crop Evolution*, v. 51, p. 205-209, 2004.

SUWARNO, W.B.; PIXLEY, K.V.; PALACIOS-ROJAS, N.; KAEPLER, S.M.; BABU, R. Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. *Theoretical and Applied Genetics*, v. 128, p. 851-864, 2015.

TAHERI, S.; ABDULLAH, T.L.; YUSOP, M.R.; HANAFI, M.M.; SAHEBI, M.; AZIZI, P.; SHAMSHIRI, R.R. Mining and development of novel SSR markers using next generation sequencing (NGS) data in plants. *Molecules*, v. 23, p. 399, 2018.

TAKAHAMA, U.; ONIKI, T. Effects of ascorbate on the oxidation of derivatives of hydroxycinnamic acid and the mechanism of oxidation of sinapic acid by cell wall-bound peroxidases. *Plant Cell Physiology*, v. 35, p. 593-600, 1994.

TAYLOR, S.H.; HULME, S.P.; REES, M.; RIPLEY, B.S.; WOODWARD, F.I.; OSBORNE, C.P. Ecophysiological traits in C3 and C4 grasses: a phylogenetically controlled screening experiment. *New Phytologist*, v. 185, n. 3, p. 780–791, 2010.

TEHRANI, M.S.; MARDI, M.; SAHEBI, J.; CATALÁN, P.; DÍAZ-PÉREZ, A. Genetic diversity and structure among Iranian tall fescue populations based on genomic-SSR and EST-SSR marker analysis. *Plant Systematics and Evolution*, v. 282, p. 57-70, 2009.

TETREAU, H.M.; GRIES, T.; PALMER, N.A.; FUNNELL-HARRIS, D.; SATO, S.; GE, Z.; SARATH, G.; SATTLER, S.E. Overexpression of ferulate 5-hydroxylase increases syringyl units in *Sorghum bicolor*. *Plant Molecular Biology*, v. 103, n. 3, p. 269-285, 2020.

THE UNIPROT CONSORTIUM. UniProt: a worldwide hub of protein knowledge in 2021. *Nucleic Acids Research*, v. 49, 2021.

TIAN, H.; YANG, Y.; WANG, R.; FAN, Y.; YI, H.; JIANG, B.; WANG, L.; REN, J.; XU, L.; ZHANG, Y.; GE, J.; LIU, Y.; WANG, F.; ZHAO, J. Screening of 200 Core SNPs and the Construction of a Systematic SNP-DNA Standard Fingerprint Database with More Than 20,000 Maize Varieties. *Agriculture*, v. 11, p. 597, 2021.

TORRES GONZÁLEZ, A.M.; MORTON, C.M. Molecular and morphological phylogenetic analysis of *Brachiaria* e *Urochloa* (Poaceae). *Molecular Phylogenetics and Evolution*, v. 37, p. 36-44, 2005.

TRAPNELL, C.; WILLIAMS, B.A.; PERTEA, G.; MORTAZAVI, A.; KWAN, G.; VAN BAREN, M.J.; SALZBERG, S.L.; WOLD, B.J.; PATCHER, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, v. 28, p. 511-515, 2010.

TRIVIÑO, N.J.; PEREZ, J.G.; RECIO, M.E.; EBINA, M.; YAMANAKA, N.; TSURUTA, S.-I.; ISHITANI, M.; WORTHINGTON, M. Genetic Diversity and Population Structure of *Brachiaria* Species and Breeding Populations. *Crop Science*, v. 57, p. 2633-2644, 2017.

TROPICOS. Missouri Botanical Garden. Disponível em: <http://www.tropicos.org>. Acesso em 05 junho 2019.

TSAI, C.C.; CHEN, Y.K.H.; CHEN, C.H.; WENG, I.S.; TSAI, C.M.; LEE, S.R.; LIN, Y.S.; CHIANG, Y.C. Cultivar identification and genetic relationship of mango (*Mangifera indica*) in Taiwan using 37 SSR markers. *Scientia Horticulturae*, v. 164, p. 196-201, 2013.

UNTERGASSER, A.; CUTCUTACHE, I.; KORESSAAR, T.; YE, J.; FAIRCLOTH, B.C.; REMM, M.; ROZEN, S.G. Primer3 – new capabilities and interfaces. *Nucleic Acids Research*, v. 40, n. 15, p. e115, 2012.

VALENTINE, J.; CLIFTON-BROWN, J.; HASTINGS, A.; ROBSON, P.; ALLISON, G.; SMITH, P. Food vs. Fuel: the use of land for lignocellulosic “next generation” energy crops tha minimize competition with primary food production. *GCB Bioenergy*, v. 4, p. 1-19, 2012.

VAN ACKER, R.; VANHOLME, R.; STORME, V.; MORTIMER, J.C.; DUPREE, P.; BOERJAN, W. Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnology for Biofuels*, v.6, p. 46, 2013.

VAN DER WEIJDE, T.; KAMEI, C.L.A.; TORRES, A.F.; VERMERRIS, W.; DOLSTRA, O.; VISSER, R.G.F.; TRINDADE, L.M. The potential of C4 grasses for cellulosic biofuel production. *Frontiers in Plant Science*, v. 4, 2013.

VAN SOEST, P.J.; ROBERTSON, J.B. LEWIS, B.A. Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *Journal of Dairy Science*, v. 74, p. 3583-3597, 1991.

VAN SOEST, P.J. *Nutritional Ecology of the Ruminant*. Ithaca, NY: Cornell University Press, 1994.

VANHOLME, R.; DEMEDTS, B.; MORREEL, K.; RALPH, J.; BOERJAN, W. Lignin biosynthesis and structure. *Plant Physiology*, v. 153, p. 895-905, 2010.

VANHOLME, R.; DE MEESTER, B.; RALPH, J.; BOERJAN, W. Lignin biosynthesis and its integration into metabolism. *Current Opinion in Biotechnology*, v. 56, p. 230-239, 2019.

VELDKAMP, J.F. A revision of *Cenchrus* incl. *Pennisetum* (Gramineae) in Malesia with some general nomenclatural notes. *Blumea*, v. 59, p. 59-75, 2014.

VERLOOVE, F. New combinations in *Cenchrus* (Paniceae, Poaceae) in Europe and the Mediterranean area. *Willdenowia*, v. 42, p. 77-78, 2012.

VICENTINI, R.; BOTTCHEER, A.; BRITO, M.S.; DOS SANTOS, A.B.; CRESTE, S.; LANDELI, M.G.D.; CESARINO, I.; MAZZAFERA, P. Large-Scale Transcriptome Analysis of Two Sugarcane Genotypes Contrasting for Lignin Content. *PLoS ONE*, v. 10, n. 8, 2015.

VIEIRA, A.C. Caracterização da biomassa proveniente de resíduos agrícolas. 2012. 56 f. Dissertação (Mestrado em Energia na Agricultura) – UNIOESTE, Cascavel, 2012.

VIGNA, B.B.Z.; DE OLIVEIRA, F.A.; TOLEDO-SILVA, G.; DA SILVA, C.C.; DO VALLE, C.B.; DE SOUZA, A.P. Leaf transcriptome of two highly divergent genotypes of *Urochloa humidicola* (Poaceae), a tropical polyploid forage grass adapted to acidic soils and temporary flooding areas. *BMC Genomics*, v. 17, p. 910, 2016a.

VIGNA, B.B.Z.; SANTOS, J.C.S.; JUNGSMANN, L.; DO VALLE, C.B.; MOLLINARI, M.; PASTINA, M.M.; PAGLIARINI, M.S.; GARCIA, A.A.F.; SOUZA, A.P. Evidence of Allopolyploid in *Urochloa humidicola* Based on Cytological Analysis and Genetic Linkage Mapping. *PLoS ONE*, v.11, n. 4, 2016b.

VIJAY, N. POELSTRA, J.W.; KÜNSTNER, A.; WOLF, J.B.W. Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Molecular Ecology*, v. 22, p. 620-634, 2013.

VOGT, T. Phenylpropanoid biosynthesis. *Molecular Plant*, v. 3, p. 2-20, 2010.

XIE, R.J.; ZHOU, J.; WANG, G.Y.; ZHANG, S.M.; CHEN, L.; GAO, Z.S. Cultivar identification and genetic diversity of Chinese bayberry (*Myrica rubra*) accessions based on fluorescent SSR markers. *Plant Molecular Biology Reporter*, v. 29, p. 554-562, 2011a.

XIE, W.; ZHANG, X.; CAI, H.; HUANG, L.; PENG, Y.; MA, X. Genetic maps of SSR and SRAP markers in diploid orchardgrass (*Dactylis glomerata* L) using the pseudo-testcross strategy. *Genome*, v. 54, n. 3, p. 212-221, 2011b.

XU, Z.; ZHANG, D.; HU, J.; ZHOU, X.; YE, X.; REICHEL, K.L.; STEWART, N.R.; SYRENNE, R.D.; YANG, X.; GAO, P.; SHI, W.; DOEPPKE, C.; SYKES, R.W.; BURRIS, J.N.; BOZELL, J.J.; CHENG, Z.M.; HAYES, D.G.; LABBE, N.; DAVIS, M.; STEWART JUNIOR, C.N.; YUAN, J.S. Comparative genome analysis of lignina biosynthesis gene families across the plant kingdom. *BMC Bioinformatics*, 2009.

XU, W.; TANG, W.; WANG, C.; GE, L.; SUN, J.; QI, X.; HE, Z.; ZHOU, Y.; CHEN, J.; XU, Z.; MA, Y-Z.; CHEN, M. *SiMYB56* Confers Drought Stress Tolerance in Transgenic Rice by Regulating Liginin Biosynthesis and ABA Signaling Pathway. *Frontiers in Plant Science*, v. 11, p. 785, 2020.

XU, X.; CROW, M.; RICE, B.R.; LI, F.; HARRIS, B.; LIU, L.; DEMESA-AREVALO, E.; LU, Z.; WANG, L.; FOX, N.; WANG, X.; DRENKOW, J.; LUO, A.; CHAR, S.N.; YANG, B.; SYLVESTER, A.W.; GINGERAS, T.R.; SCMITZ, R.J.; WARE, D.; LIPKA, A.E.; GILLIS, J.; JACKSON, D. Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Developmental Cell*, v. 56, n. 4, p. 557-568, 2021.

WAGHORN, G.; CLARK, D. Feeding value of pastures for ruminants. *New Zealand Veterinary Journal*, v. 52, p. 332-341, 2004.

WAGNER, G.P.; KIN, K.; LYNCH, V.J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*, v. 131, p. 281-285, 2012.

WAI, C.M.; ZHANG, J.; JONES, T.C.; NAGAI, C.; MING, R. Cell wall metabolism and hexose allocation contribute to biomass accumulation in high yielding extreme segregants of a *Saccharum* interspecific F2 population. *BMC Genomics*, v. 18, p. 773, 2017.

WALSH, B.M.; HOOT, S.B. Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: the chloroplast atpB-rbcL spacer region and nuclear waxy introns. *International Journal of Plant Sciences*, v. 162, p. 1409-1418, 2001.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews*, v. 10, p. 57-63, 2009.

WANG, F.G.; TIAN, H.L.; ZHAO, J.R.; YI, H.M.; WANG, L.; SONG, W. Development and characterization of a core set of SSR markers for fingerprint analysis of Chinese maize varieties. *Maydica*, v. 56, n. 1, p. 1-11, 2011.

WANG, H.Z.; DIXON, R.A. On-off switches for secondary cell wall biosynthesis. *Molecular Plant*, v. 5, p. 297-303, 2012.

WANG, H.; CHEN, N.F.; ZHENG, J.Y.; WANG, W.C.; PEI, Y.Y.; ZHU, G.P. Isolation and characterization of eleven polymorphic microsatellite loci for the valuable medicinal plant *Dendrobium huoshanense* and cross-species amplification. *International Journal of Molecular Sciences*, v. 13, p. 16779-16784, 2012.

WATSON, L.; DALLWITZ, M.J. The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, 1992. Disponível em: <http://delta-intkey.com>. Acesso em 16 abril 2019.

WATTS, K.T.; MIJTS, B.N.; LEE, P.C.; MANNING, A.J.; SCHMIDT-DANNERT, C. Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family. *Chemistry & Biology*, v. 13, p. 1317-1326, 2006.

WEBER, A.P.M.; WEBER, K.L.; CARR, K.; WILKERSON, C.; OHLROGGUE, J.B. Sampling the arabidopsis transcriptome with massively parallel pyrosequencing. *Plant Physiology*, v. 144, p. 32-42, 2007.

WEISS, W.P.; HALL, M.B. Laboratory Methods for Evaluating Forage Quality. In: MOORE, K.J.; COLLINS, M.; NELSON, C.J.; REDFEARN, D.D. (Org.) *Forages: The Science of Grassland Agriculture*, v. 2, 7. ed. cap. 36, p. 659-672, 2020.

WENG, J.-K.; LI, X.; BONAWITZ, N.D.; CHAPPLE, C. Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Current Opinion in Biotechnology*, v. 19, p. 166-172, 2008.

WHITEHEAD, C.; GARRIDO, F.J.O.; REYMOND, M.; SIMISTER, R.; DISTELFELS, A.; ATIENZA, S.G.; PISTON, F.; GOMEZ, L.D.; MCQUEEN-MASON, S.J. A glycosyl transferase family 43 protein involved in xylan biosynthesis is associated with straw digestibility in *Brachypodium distachyon*. *New Phytologist*, v. 218, p. 974-985, 2018.

WU, J.W.; HU, C.Y.; SHAHID, M.Q.; GUO, H.B.; ZENG, Y.Z.; LIU, X.D.; LU, Y.G. Analysis on genetic diversification and heterosis in autotetraploid rice. Springer Plus, v. 2, n. 1, p. 439, 2013.

WU, D.C.; YAO, J.; HO, K.S.; LAMBOWITZ, A.M. WILKE, C.O. Limitations of alignment-free tools in total RNA-seq quantification. BMG Genomics, v. 19, p. 510, 2018.

WULFERT, S.; KRUEGER, S. *Phosphoserine Aminotransferase1* is Part of the Phosphorylated Pathways for Serine Biosynthesis and Essential for Light and Sugar-Dependent Growth Promotion. Frontiers in Plant Science, v. 9, p. 1712, 2018.

YAN, Q.; LI, J.; LU, L.; GAO, L.; LAI, D.; YAO, N.; YI, X.; WU, Z.; LAI, Z.; ZHANG, J. Integrated analyses of phenotype, phytohormone, and transcriptome to elucidate the mechanism governing internode elongation in two contrasting elephant grass (*Cenchrus purpureus*) cultivars. Industrial Crops & Products, v. 170, 2021a.

YAN, Q.; WU, F.; XU, P.; SUN, Z.; LI, J.; GAO, L.; LU, L.; CHEN, D.; MUKTAR, M.; JONES, C.; YI, X.; ZHANG, J. The elephant grass (*Cenchrus purpureus*) genome provides insights into anthocyanidin accumulation and fast growth. Molecular Ecology Resources, v. 21, n. 2, 2021b.

YASODHA, R.; VASUDEVA, R.; BALAKRISHNAN, S.; SAKTHI, A.R.; ABEL, N.; BINAI, N.; RAJASHEKAR, B.; BACHPAI, V.K.W.; PILLAI, C.; DEV, S.A. Draft genome of a high value tropical timber tree, Teak (*Tectona grandis* L. f): insights into SSR diversity, phylogeny and conservation. DNA Research, v. 25, n. 4, p. 409-419, 2018.

YASUDA, M.; ISHII, V.; OHTA, K. Napier Grass (*Pennisetum purpureum* Schumach) as Raw Material for Bioethanol Production: Pretreatment, Saccharification, and Fermentation. Biotechnology and Bioprocess Engineering, v. 19, p. 943-950, 2014.

YE, Q.; LIU, X.; BIAN, W.; ZHANG, Z.; ZHANG, H. Over-expression of transcription factor ARK1 gene leads to down-regulation of lignin synthesis related genes in hybrid poplar '717'. Scientific Reports, v. 10, n. 1, p. 8549, 2020.

YI, L.; LIU, L.; MELSTED, P.; PATCHER, L. A direct comparison of genome alignment and transcriptome pseudoalignment. bioRxiv, 2018.

YIN, M.; MA, H.; WANG, M.; CHU, G.; LIU, Y.; XU, C.; ZHANG, X.; WANG, D.; CHEN, S. Transcriptome analysis of flowering regulation by sowing date in *Japonica* Rice (*Oryza sativa* L.). Scientific Reports, v. 11, 2021.

YU, M.; WANG, M.; GYALPO, T.; BASANG, Y. Stem lodging resistance in hulless barley: Transcriptome and metabolome analysis of lignin biosynthesis pathways in contrasting genotypes. *Genomics*, v. 113, p. 935-943, 2020.

YUAN, J.S.; TILLER, K.H.; AL-AHMAD, H.; STEWART, N.R.; STEWART JUNIOR, C.N. Plants to power: bioenergy to fuel the future. *Trends in Plant Science*, v. 13, p. 421-429, 2014.

ZAINAL-ABIDIN, R.A.; ZAINAL, Z.; MOHAMED-HUSSEIN, Z.A.; ABU-BAKAR, N.; RAZAK, M.S.F.A.; SIMOH, S.; SEW, Y.S. RNA-seq data from whole rice grains of pigmented and non-pigmented Malaysian rice varieties. *Data in Brief*, v. 30, p. 105432, 2020.

ZARGAR, M.; ROMANOVA, E.; SHMELKOVA, E.; TRIFONOVA, A.; KEZIMANA, P. AFLP-analysis of genetic diversity in soybean [*Glycine max* (L.) Merr.] cultivars Russian and foreign selection. *Agronomy Research*, v. 15, n. 5, p. 2217-2225, 2017.

ZENG, X.; SHENG, J.; ZHU, F.; WEI, T.; ZHAO, L.; HU, X.; ZHENG, X.; ZHOU, F.; HU, Z.; DIAO, Y.; JIN, S. Genetic, transcriptional, and regulatory landscape of monoglignol biosynthesis pathway in *Miscanthus x giganteus*. *Biotechnology for Biofuels*, v. 13, p. 179, 2020.

ZHAI, S.H.; YIN, G.S.; YANG, X.H. Population Genetics of the Endangered and Wild Edible Plant *Ottelia acuminata* in Southwestern China Using Novel SSR Markers. *Biochemical Genetics*, v. 56, p. 235-254, 2018.

ZHANG, W.; YUAN, Q.; JIANG, D.; ZHANG, Y.; HUANG, L. Development and characterization of microsatellite markers for the medicinal plant *Scutellaria baicalensis* (Lamiaceae). *Biochemical Systematics and Ecology*, v. 54, p. 267-271, 2014.

ZHANG, J.; LI, M.; BRYAN, A.C.; YOO, C.G.; ROTTMANN, W.; WINKELER, K.A.; COLLINS, C.M.; SINGAN, V.; LINDQUIST, E.A.; JAWDY, S.S.; GUNTER, L.E.; ENGLE, N.L.; YANG, X.; BARRY, K.; TSCHAPLINSKI, T.J.; SCHMUTZ, J.; PU, Y.; RAGAUSKAS, A.J.; TUSKAN, G.A.; MUCHERO, W.; CHEN, J.G. Overexpression of a serine hydroxymethyltransferase increases biomass production and reduces recalcitrance in the bioenergy crop *Populus*. *Sustainable Energy & Fuels*, v. 3, p. 195, 2019.

ZHANG, W.; ZHANG, S.; LU, X.; LIU, X.; DONG, G.; XIA, T. Tissue-specific Transcriptome analysis reveals lignocellulose synthesis regulation in elephant grass (*Pennisetum purpureum* Schum). BMC Plant Biology, v. 20, p.528, 2020.

ZHAO, Q.; DIXON, R.A. Transcriptional networks for lignin biosynthesis: more complex than we thought? Trends in Plant Science, v. 16, p. 227-233, 2011.

ZHONG, R.; YE, Z.H. Transcriptional regulation of lignin biosynthesis. Plant Signaling & Behavior, v. 4, n. 11, p. 1028-1034, 2009.

ZHOU, S.; WANG, C.; FRAZIER, T.P.; YAN, H.; CHEN, P.; CHEN, Z.; HUANG, L.; ZHANG, X.; PENG, Y.; MA, X.; YAN, Y. The first Illumina-based *de novo* transcriptome analysis and molecular marker development in Napier grass (*Pennisetum purpureum*). Molecular Breeding, v. 38, p. 95, 2018.