

Universidade Federal de Minas Gerais

Instituto de Ciências Exatas

Departamento de Química

Jéssika Thayanne da Silva

APLICAÇÃO DA METABOLÔMICA POR RMN NA DIFERENCIAÇÃO DE  
CULTIVARES DE SORGO BUSCANDO UM AUMENTO NA EFICIÊNCIA DA  
PRODUÇÃO DE ETANOL

Belo Horizonte

2024

UFMG/ICEx/DQ. 1.680

D. 873

Jéssika Thayanne da Silva

APLICAÇÃO DA METABOLÔMICA POR RMN NA DIFERENCIAÇÃO DE  
CULTIVARES DE SORGO BUSCANDO UM AUMENTO NA EFICIÊNCIA DA  
PRODUÇÃO DE ETANOL

Dissertação apresentada ao Departamento de Química  
do Instituto de Ciências Exatas da Universidade  
Federal de Minas Gerais, como requisito parcial para  
a obtenção do grau de Mestre em Química.

Orientadora: Prof<sup>a</sup> Dra<sup>a</sup> Lúcia Pinheiro Santos Pimenta

Belo Horizonte

2024

Ficha Catalográfica

S586a Silva, Jéssika Thyanne da.  
2024 Aplicação da metabolômica por RMN na diferenciação de cultivares de sorgo  
D buscando um aumento na eficiência da produção de etanol [manuscrito] / Jéssika  
Thyanne da Silva. 2024.  
85 f. : il., gráfs., tabs.

Orientadora: Lúcia Pinheiro Santos Pimenta.

Dissertação (mestrado) – Universidade Federal de Minas Gerais –  
Departamento de Química.

Bibliografia: f. 75-81.

Apêndices: f. 82-85.

1. Química orgânica – Teses. 2. Quimiometria – Teses. 3. Metabolômica –  
Teses. 4. Sorgo – Teses. 5. Biocombustíveis – Teses. 6. Bioetanol – Teses. 7.  
Biomassa – Teses. 8. Ressonância magnética nuclear – Teses. I. Pimenta, Lúcia  
Pinheiro Santos, Orientadora. II. Título.

CDU 043



UNIVERSIDADE FEDERAL DE MINAS GERAIS



## "Aplicação da Metabolômica por RMN na Diferenciação de Cultivares de Sorgo Buscando um Aumento na Eficiência da Produção de Etanol"

**Jéssika Thayanne da Silva**

Dissertação aprovada pela banca examinadora constituída pelos Professores:

Profª. Lúcia Pinheiro Santos Pimenta - Orientadora  
UFMG

Prof. Etelvino Henrique Novotny  
EMBRAPA

Profª. Vânia Márcia Duarte Pasa  
UFMG

Belo Horizonte, 29 de maio de 2024.



Documento assinado eletronicamente por **Etelvino Henrique Novotny, Usuário Externo**, em 29/05/2024, às 16:28, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lucia Pinheiro Santos Pimenta, Professora do Magistério Superior**, em 04/06/2024, às 10:28, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Vanya Marcia Duarte Pasa, Professora do Magistério Superior**, em 29/07/2024, às 17:11, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3271621** e o código CRC **2C091577**.

## **AGRADECIMENTOS**

Ao meu alicerce, no qual me inspiro e onde me moldei como sou. Por isso, agradeço aos meus pais João Bosco e Ivaneide e aos meus irmãos Tatianne e João Paulo.

À Bru, meu apoio e incentivo diário, e aos meus amigos: da graduação, do ACS Student Chapter e da UFMG como um todo, assim como os que vieram antes e de fora do campus. Obrigada por caminharem comigo.

À Universidade Federal de Minas Gerais e ao Departamento de Química, pela valiosa oportunidade de realizar meu Mestrado em uma instituição tão renomada.

À professora Dr<sup>a</sup> Lúcia Pinheiro, pelos ricos ensinamentos, pela companhia, conselhos e por toda a paciência durante essa etapa.

Ao Dr. Alan Rodrigues Teixeira Machado, por toda a ajuda e colaboração, principalmente quando me encontrei perdida com meus experimentos de RMN.

Ao Laboratório de Ensaio de Combustíveis – LEC, e ao Programa de Recursos Humanos da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – PRH-ANP, suportado com recursos provenientes do investimento de empresas petrolíferas qualificadas na Cláusula de PD&I da Resolução ANP nº 918, de 10/03/2023, pela concessão da bolsa de estudos e todo o financiamento.

Ao Núcleo de Extensão e Prestação de Serviço (NEPS) do Departamento de Química da UFMG, pela infraestrutura, em especial ao Laboratório de Ressonância Magnética de Alta Resolução (LAREMAR).

À Embrapa Milho e Sorgo, pela parceria durante a realização do projeto de pesquisa, e aos que lá gentilmente me acolheram, em especial à Dr<sup>a</sup> Maria Lúcia Ferreira Simeone, à Cristiane de Carvalho Guimarães e ao Rafael de Araújo Miguel.

*“Eu adorava as descobertas que a química oferecia.  
Adorava a emoção de fazer experimentos, o desafio do  
ensaio e erro. Gostava do caráter intrigante daquilo. Há na  
química algo de primitivo, algo que desafia a explicação.  
Ou se sente isso, ou não”*

Patrick Rothfuss

## RESUMO

*Sorghum bicolor* L. Moench é uma excelente alternativa à cana-de-açúcar e ao milho na produção de bioetanol uma vez que é uma cultura mais barata, possui alto teor de amido, alto potencial de rendimento, baixa demanda de insumos e relativa tolerância à seca. Quatro genótipos de sorgo, duas linhagens (SC084, CMSXS180) e dois híbridos (BRS305, BR501) foram avaliados quanto ao seu perfil metabólico e eficiência na produção de etanol. A análise metabolômica tem crescido como uma ferramenta utilizada na compreensão de como fatores genéticos e ambientais afetam as variações fenotípicas de plantas. Neste estudo foi avaliada a utilização da metabolômica por RMN na diferenciação de quatro genótipos de sorgo e a correlação desses perfis metabólicos com a melhor eficiência na produção de etanol. Foram determinados os teores de compostos fenólicos totais, taninos, antocianidinas, amido total, extrato etéreo, matéria mineral, proteína bruta e a atividade antioxidante para as cultivares. A cultivar SC084 foi a que apresentou maiores teores de compostos fenólicos totais, taninos, antocianidinas e proteína bruta, e o menor teor em amido total. A análise metabolômica por RMN não objetivada mostrou que as diferentes cultivares apresentam perfis metabólicos diferentes, principalmente na região de aromáticos, e que a cultivar SC084 apresenta perfil diferenciado das demais destacando-se na região em  $\delta$  7,06-5,98. As cultivares BR501 e CMSXS180 apresentaram maiores rendimentos na produção de bioetanol pelo processo de sacarificação e fermentação simultâneas. Análises univariadas e multivariadas permitiram determinar correlações entre os metabólitos, o teor de taninos e o tratamento de desengorduramento. O teor de taninos se encontra altamente correlacionado negativamente ao teor de etanol, enquanto o tratamento dos grãos não exerce influência nesse fator. Essa análise identificou os sinais de deslocamento químico condizentes com a presença de unidades flavan-3-ol presentes em proantocianidinas na composição da cultivar SC084, que apresentou menor rendimento na produção de etanol. Verificou-se também que a durrina e produtos de seu catabolismo estão presentes no metaboloma da linhagem CMSXS180 diferenciando-a das demais. Aplicação do método de classificação supervisionado floresta aleatória (*random forest*) mostrou que sinais característicos de unidades flavan-3-ol são importantes na classificação de amostras com alto teor de taninos, sugerindo a possibilidade da geração de um modelo preditivo para o rendimento em etanol baseado na intensidade dos sinais de RMN de  $^1\text{H}$ . A metabolômica por RMN se apresentou como um método rápido na diferenciação metabólica das cultivares de sorgo e que se presta ao estudo da correlação dos perfis metabólicos ao processo de obtenção de etanol. Este estudo representa um ponto de partida para a aplicação desta ferramenta na cadeia produtiva do sorgo visando sua utilização na matriz energética.

Palavras-chave: *Sorghum bicolor*; bioetanol; quimiometria; metabolômica por RMN de  $^1\text{H}$ ; bioenergia; biomassa.

## ABSTRACT

*Sorghum bicolor* L. Moench is an excellent alternative to sugar cane and maize in the production of bioethanol since it is a cheaper crop, has a high starch content, high yield potential, low input demand, and relative tolerance to drought. Four sorghum genotypes, two lines (SC084, CMSXS180), and two hybrids (BRS305, BR501) were evaluated for their metabolic profile and efficiency in ethanol production. Metabolomic analysis has grown as a tool to understand how genetic and environmental factors affect phenotypic variations of plants. In this study, the use of NMR-based metabolomics was evaluated in differentiating four sorghum genotypes and its application in correlating these metabolic profiles with better efficiency in ethanol production. The content of total phenolic compounds, tannins, anthocyanins, total starch, ether extract, mineral matter, crude protein, and antioxidant activity were determined for the cultivars. Cultivar SC084 was the one with the highest levels of total phenolic compounds, tannins, anthocyanins, and crude protein, and the lowest total starch content. Non-targeted NMR-based metabolomic analysis showed that the different cultivars have different metabolic profiles, mainly in the aromatic region, and that the SC084 cultivar has a different profile from the others, standing out in the region at  $\delta$  7.06-5.98. BR501 and CMSXS180 cultivars showed higher yields in bioethanol production through the simultaneous saccharification and fermentation process. Correlations between metabolites, tannin content, and degreasing treatment were possible through univariate and multivariate analyses. Tannin content is highly negatively correlated with ethanol content, while treatment of the grains does not influence this factor. This analysis identified the chemical shifts consistent with the presence of flavan-3-ol units from proanthocyanidins and the ones observed in the metabolic profile of the SC084 cultivar, which presented a lower yield in ethanol production. It was found that the presence of durrin and its catabolism products are present in the metabolome of the line CMSXS180, differentiating it from the others. The supervised random forest classification method showed that characteristic signals of flavan-3-ol units are important in classifying samples with high tannin content, suggesting the possibility of a predictive model for ethanol yield based on the intensity of  $^1\text{H}$  NMR signals. NMR-based metabolomics presented itself as a quick method for the metabolic differentiation of sorghum cultivars and it lends itself to studying the correlation of metabolic profiles with the process of obtaining ethanol. This study represents a starting point for the application of this tool in the sorghum production chain aiming for its use in the energy matrix.

Keywords: *Sorghum bicolor*; bioethanol; chemometrics;  $^1\text{H}$  NMR-based metabolomics; bioenergy; biomass.

## LISTA DE FIGURAS

- Figura 1. Acoplamentos possíveis que ocorrem no COSY: geminais (Hc-Hd no carbono C4) e vicinais (Hb-Hc nos carbonos C3 e C4). A correlação resultante do acoplamento Ha-Hb dos carbonos C1 e C3 não ocorre.....26
- Figura 2. (a) Acoplamento direto observado pelo HSQC; e (b) possibilidades de acoplamento à longa distância observado nos experimentos de HMBC. ....27
- Figura 3. Amostras de sorgo granífero. ....30
- Figura 4. Equação representativa da reação da vanilina com um tanino condensado..... 33
- Figura 5. Estrutura das antocianidinas identificadas nos grãos de sorgo .....35
- Figura 6. Balões fechados com rolhas de silicone acopladas a airlocks em incubadora shaker preparados para o processo de sacarificação e fermentação simultâneas. ....41
- Figura 7. Teor de compostos fenólicos totais determinados por duas metodologias diferentes para as amostras de sorgo *in natura*. GAE (Ácido Gálico Equivalente). ....45
- Figura 8. Teor de amido total (%) para as amostras de sorgo *in natura* e desengorduradas. ...47
- Figura 9. Perfil do espectro no ultravioleta-visível (200 a 600 nm) para as amostras de sorgo, indicando os comprimentos de onda máximos de cada amostra. ....51
- Figura 10. a) Espectros de RMN de  $^1\text{H}$  dos extratos dos grãos de sorgo *in natura* e desengordurados (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0). ....53
- Figura 11. Gráfico de scores das componentes principais PC1 (51,7%) e PC2 (32,5%) calculadas a partir dos dados espectrais de RMN de  $^1\text{H}$  das amostras de sorgo *in natura* e desengordurados, construído com o auxílio do *software* MetaboAnalyst 6.0. ....54
- Figura 12. Gráfico de loadings das componentes PC1 e PC2, obtido a partir dos dados espectrais de RMN de  $^1\text{H}$  das amostras de sorgo *in natura* e desengordurados.....55

|   |    |
|---|----|
| Figura 13. Espectro de RMN de $^1\text{H}$ da amostra SC084 in natura (600 MHz, Metanol- $d_4$ e tampão $\text{KH}_2\text{PO}_4$ em $\text{D}_2\text{O}$ com 0,01% TSP- $d_4$ (m/m), pH 6,0). Expansão em $\delta$ 5,90-7,10 ppm.<br>.....  | 56 |
| Figura 14. Mapa de contornos COSY (A) e Espectro $J$ -resolvido (B) do grão de sorgo da cultivar SC084 (600 MHz, metanol- $d_4$ e tampão $\text{KH}_2\text{PO}_4$ em $\text{D}_2\text{O}$ com 0,01% TSP- $d_4$ (m/m), pH 6,0) com expansão em $\delta$ 5,80-7,20. ....  | 58 |
| Figura 15. Análise Hierárquica de Clusters (HCA) dos dados de RMN de $^1\text{H}$ das amostras de sorgo construído com o auxílio do <i>software</i> MetaboAnalyst. ....   | 60 |
| Figura 16. Mapa de calor dos dados espectrais de RMN de $^1\text{H}$ das amostras de sorgo in natura e desengorduradas em estudo construído com o auxílio do <i>software</i> MetaboAnalyst, onde as linhas representam os sinais de ressonância (ppm) e as colunas representam as amostras agrupadas de acordo com a Análise Hierárquica de Clusters (HCA) dos dados. Valores em vermelho com cores mais escuras indicam maior importância dos valores e cores em azul indicam valores menos importantes..... | 61 |
| Figura 17. Estruturas das moléculas de amido e glicose, indicando as ligações $\alpha$ -1,4 e $\alpha$ -1,6 presentes na molécula de amido.....   | 63 |
| Figura 18. Teores de etanol, em $\text{mL g}^{-1}$ de sorgo, obtidos a partir da fermentação das amostras <i>in natura</i> e desengorduradas. ....  | 64 |
| Figura 19. Diagrama das correlações entre os metadados tratamento (grãos <i>in natura</i> e desengordurados), teor de taninos (alto, intermediário e baixo) e teor de etanol. ...   | 66 |
| Figura 20. Os 25 principais <i>bins</i> correlacionados entre as variáveis ( <i>bins</i> ), os metadado de interesse (teor de etanol) e a covariável de interesse (teor de taninos). Em azul, deslocamentos químicos que afetam negativamente o teor de etanol. Em rosa, os que afetam positivamente. ....  | 67 |
| Figura 21. Espectro de RMN de $^1\text{H}$ das amostras CMSXS180 (A) e BR501 (B) com expansão na região entre $\delta$ 5,0-10,0 (600 MHz, Metanol- $d_4$ e tampão $\text{KH}_2\text{PO}_4$ em $\text{D}_2\text{O}$ com 0,01% TSP- $d_4$ (m/m), pH 6,0). ....  | 69 |

- Figura 22. A) Mapa de contornos *J*-resolvido expansão em  $\delta$  6,0-8,6; B) Mapa de contornos COSY para o genótipo de sorgo CMSX180 *in natura* com a expansão entre  $\delta$  6,0 e 8,1 (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0). ..... 69
- Figura 23. Mapa de contornos HSQC da cultivar CMSXS180 *in natura* destacando sinais atribuídos ao ácido *p*-glicosiloxy-2-hidroxifenilacético ( $\delta$  7,39 e 7,10), à durrina ( $\delta$  7,29 e 6,84) e ao ácido *p*-hidroxibenzoico ( $\delta$  7,81 e 6,86). (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0). ..... 70
- Figura 24. A) Modelo de classificação *random forest* (RF) mostrando o erro medido para cada classe usando o algoritmo de aprendizado de máquina, B) *Bins* classificados por suas contribuições para a precisão da classificação (diminuição média na precisão). ..... 72
- Figura 25. Espectros de RMN de  $^1\text{H}$  da amostra de sorgo BRS305 *in natura* (verde) e desengordurada (marrom). 600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0..... 83
- Figura 26. A) Mapa de contornos HSQC da cultivar SC084 *in natura* destacando a região do anel A de unidades flavonoídicas e sinais característicos de ácidos graxos insaturados (AGI); B) expansão entre  $\delta$  5 e 7,5 destacando a região de sinais dos anéis A e B de unidades flavonoídicas. (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0). ..... 84

## LISTA DE TABELAS

|  |    |
|--|----|
| Tabela 1. Descrição das amostras de <i>Sorghum bicolor</i> L. Moench granífero cedidas pela Embrapa Milho e Sorgo para desenvolvimento do projeto de pesquisa .....                          | 30 |
| Tabela 2. Perda de massa (%) das amostras de sorgo ocasionadas pelo tratamento de desengorduramento dos grãos e teor de extrato hexânico (% m/m).....  | 44 |
| Tabela 3. Teor de taninos para as amostras de sorgo in natura .....  | 46 |
| Tabela 4. Teor de antocianidinas, em $\mu\text{g g}^{-1}$ , determinado por CLAE.....  | 48 |
| Tabela 5. Composição centesimal das amostras de sorgo in natura e desengorduradas (% m/m) para as diferentes cultivares .....  | 49 |
| Tabela 6. Atividade antioxidante das amostras de sorgo in natura, expressa em equivalentes de Trolox ( $\mu\text{mol TE g}^{-1}$ amostra) .....  | 51 |
| Tabela 7. Atribuições de metabólitos identificados no sorgo pelo perfil no RMN, com deslocamentos químicos ( $\delta$ , ppm), padrão de divisão e constantes de acoplamento ( $J$ , Hz)..... | 59 |
| Tabela 8. Rendimento do processo de fermentação para as amostras de sorgo in natura e desengorduradas .....  | 65 |
| Tabela 9. Teor de compostos fenólicos totais.....  | 82 |
| Tabela 10. Teores de glicose, glicerol e etanol para cada grama de amostra de sorgo in natura e desengorduradas após o processo de fermentação .....   | 82 |

## LISTA DE SIGLAS E ABREVIATURAS

\_d – desengordurada (sinaliza as amostras que passaram pelo processo de desengorduramento)

1D – Unidimensional

2D – Bidimensional

ABTS – Ácido 2,2'-azino-bis(3-etilbenzotiazolina-6-sulfônico)

BNDES – Banco Nacional do Desenvolvimento

CG – Cromatografia Gasosa

CGEE – Centro de Gestão e Estudos Estratégicos

CL – Cromatografia Líquida

CLAE – Cromatografia líquida de alta eficiência

COSY – Correlação Homonuclear, do inglês “*Correlation Spectroscopy*”

d – duplete

Da – Dalton (unidade de massa)

DAD – Detector por arranjo de diodos, do inglês “*Diode Array Detector*”

DDGS – Grãos Secos de Destilaria com Solúveis, do inglês “*Dried Distillers Grains with Solubles*”

DNA – Ácido desoxiribonucleico, do inglês “*Desoxyribonucleic Acid*”

EM – Espectrometria de Massas

EPE – Empresa de Pesquisa Energética

GOPOD – Ácido *p*-hidroxibenzoico e azida de sódio

HCA – Agrupamento Hierárquico de Clusters, do inglês “*Hierarchical Cluster Analysis*”

HMBC – Correlação Heteronuclear de Múltiplas Ligações, do inglês “*Heteronuclear multiplebond correlation*”

HSQC – Correlação Heteronuclear de Simples Quantum, do inglês “*Heteronuclear single quantum coherence*”

IV – Espectroscopia no Infravermelho

$J$  – Constante de acoplamento escalar spin-spin nuclear

LQ – Limite de Quantificação

m – multiplete

MDA – diminuição média na precisão, do inglês “*Mean Decrease Accuracy*”

ODS – Objetivos de Desenvolvimento Sustentável

ONU – Organização das Nações Unidas

PC – Componente Principal, do inglês “*Principal Component*”

PCA – Análise de Componentes Principais, do inglês “*Principal Component Analysis*”

PLS – Regressão por Mínimos Quadrados Parciais, do inglês “*Partial least squares regression*”

RF – Floresta aleatória, do inglês “*Random Forest*”

RFA – Associação de Combustíveis Renováveis, do inglês “*Renewable Fuel Association*”

RMN – Ressonância Magnética Nuclear

RMN-q – Ressonância Magnética Nuclear quantitativa

RNA – Ácido Ribonucleico, do inglês “*Ribonucleic Acid*”

RNA<sub>m</sub> – Ácido ribonucleico mensageiro, do inglês “*messenger Ribonucleic Acid*”

SSF - Sacarificação e Fermentação Simultâneas, do inglês “*Simultaneous Saccharification and Fermentation*”

t – tripleto

TOCSY – Correlação Total, do inglês “*Total Correlation Spectroscopy*”

TSP- $d_4$  – Sal de sódio do ácido 3-(trimetilsilil)propiónico-2,2,3,3- $d_4$

UV-Vis – Espectroscopia ultravioleta-visível

## SUMÁRIO

|   |    |
|---|----|
| 1. INTRODUÇÃO.....  | 16 |
| 2. OBJETIVOS.....   | 20 |
| 3. REFERENCIAL TEÓRICO.....   | 21 |
| 3.1 Metabolômica.....   | 21 |
| 3.2 Técnicas analíticas.....  | 23 |
| 3.2.1 Ressonância Magnética Nuclear.....                            | 23 |
| 3.2.1.1 RMN de <sup>1</sup> H.....                                  | 24 |
| 3.2.1.2 RMN bidimensional.....                                      | 25 |
| 3.2.1.2.1 <i>J</i> -resolvido.....                                  | 25 |
| 3.2.1.2.2 COSY.....   | 25 |
| 3.2.1.2.3 HSQC e HMBC.....  | 26 |
| 3.3 Análise multivariada de dados.....                              | 27 |
| 3.3.1 Análise de Componentes Principais (PCA).....                  | 28 |
| 3.3.2 Agrupamento Hierárquico de Clusters (HCA).....                | 29 |
| 4. METODOLOGIA.....   | 30 |
| 4.1 Amostras utilizadas.....  | 30 |
| 4.2 Tratamento das amostras: desengorduramento dos grãos.....       | 31 |
| 4.3 Caracterização química das amostras de sorgo.....               | 31 |
| 4.3.1 Determinação do teor de compostos fenólicos totais.....       | 31 |
| 4.3.1.1 Método 1.....   | 31 |
| 4.3.1.2 Método 2.....   | 32 |
| 4.3.2 Determinação do teor de taninos pelo método Vanilina-HCl..... | 33 |
| 4.3.3 Determinação do teor de amido total.....                      | 34 |
| 4.3.4 Determinação do perfil de antocianidinas por CLAE.....        | 35 |
| 4.3.5 Determinação da composição centesimal.....                    | 36 |
| 4.4 Caracterizações espectroscópicas do sorgo.....                  | 36 |
| 4.4.1 Determinação do perfil no ultravioleta.....                   | 36 |
| 4.4.2 Determinação da atividade antioxidante.....                   | 36 |
| 4.4.3 Determinação do perfil metabólico por RMN.....                | 38 |
| 4.4.3.1 Preparo das amostras.....                                   | 38 |
| 4.4.3.2 Obtenção dos espectros.....                                 | 38 |
| 4.5 Obtenção do etanol de sorgo.....                                | 39 |
| 4.5.1 Preparo das amostras.....                                     | 39 |
| 4.5.2 Preparo do mosto.....   | 40 |

|  |    |
|--|----|
| 4.5.3 Liquefação .....   | 40 |
| 4.5.4 Sacarificação e fermentação simultâneas .....                  | 40 |
| 4.5.5 Determinação do teor de etanol .....                           | 41 |
| 4.6 Análise multivariada dos dados de RMN de $^1\text{H}$ .....      | 42 |
| <br>   |    |
| 5. RESULTADOS E DISCUSSÃO .....                                      | 44 |
| 5.1 Tratamento das amostras: desengorduramento dos grãos .....       | 44 |
| 5.2 Caracterização química das amostras de sorgo .....               | 44 |
| 5.2.1 Determinação do teor de compostos fenólicos totais .....       | 44 |
| 5.2.2 Determinação do teor de taninos pelo método Vanilina-HCl ..... | 46 |
| 5.2.3 Determinação do teor de amido total .....                      | 47 |
| 5.2.4 Determinação do perfil de antocianidinas por CLAE .....        | 48 |
| 5.2.6 Determinação da composição centesimal .....                    | 49 |
| 5.3 Caracterizações espectroscópicas .....                           | 50 |
| 5.3.1 Determinação do perfil no ultravioleta .....                   | 50 |
| 5.3.2 Determinação da atividade antioxidante .....                   | 51 |
| 5.4 Determinação do perfil metabólico por RMN .....                  | 52 |
| 5.5 Obtenção do etanol .....   | 62 |
| <br>   |    |
| 6. CONCLUSÃO .....   | 73 |
| <br>   |    |
| REFERÊNCIAS .....  | 75 |
| <br>   |    |
| APÊNDICE A – Tabelas .....   | 82 |
| APÊNDICE B – Espectros de RMN .....                                  | 83 |
| APÊNDICE C – Cálculos .....  | 85 |

## 1. INTRODUÇÃO

A Organização das Nações Unidas coordenou, em 2016, a implementação da Agenda 2030 para o Desenvolvimento Sustentável, dentro da qual foram estabelecidos 17 Objetivos de Desenvolvimento Sustentável (ODS). Dentre os objetivos estabelecidos, o sétimo propõe “Garantir o acesso a fontes de energia fiáveis, sustentáveis e modernas para todos”, tendo como uma de suas metas “aumentar substancialmente a participação de energias renováveis na matriz energética global até o ano de 2030”. Além deste, o décimo segundo objetivo propõe “assegurar padrões de produção e de consumo sustentáveis” e traz como uma de suas metas “alcançar a gestão sustentável e o uso eficiente dos recursos naturais até 2030” (ONU, 2016).

Tendo em mente os ODS e a necessidade de se buscar fontes de energia limpa, uma boa alternativa é aumentar a matriz energética e o consumo de combustíveis de fontes renováveis como uma alternativa ao uso de combustíveis/energia fósseis. O Brasil é considerado pioneiro no consumo e produção de biocombustíveis, com destaque para o etanol produzido a partir da cana-de-açúcar (EPE, 2023), que é adicionado à gasolina como mistura obrigatória desde a década de 1930 e foi foco da criação do Programa Nacional do Álcool (Proálcool), estabelecido na década de 1970 como forma de intensificar a produção do etanol e incentivar o seu uso como combustível de forma a mitigar os efeitos da crise do petróleo à época (BNDES & CGEE, 2008; EPE, 2023; Nitsch, 1991).

A tecnologia para obtenção do etanol a partir da cana-de-açúcar já é bem estabelecida e amplamente utilizada no Brasil, sendo uma cultura de grande sucesso para a economia do país. De acordo com a Companhia Nacional de Abastecimento (Conab), a safra 2022/23 produziu 619 milhões de toneladas de cana-de-açúcar, resultando na produção de 27 bilhões de litros de etanol, além de outros subprodutos (CONAB, 2023).

No cenário mundial, o milho é a principal matéria-prima para a produção de etanol, o que acontece principalmente porque os Estados Unidos são os maiores produtores desse biocombustível, respondendo por mais da metade da produção global, utilizando o milho como a principal fonte (Manochio et al., 2017; RFA, 2024). No Brasil, a safra de milho de 2022/23 resultou na produção de 4 bilhões de litros de etanol, mostrando abertura para a diversificação das matérias-primas utilizadas na produção de etanol no país (CONAB, 2023).

A diversificação é importante porque a produção de etanol, quando focada na aplicação majoritária da cana-de-açúcar como matéria-prima, pode envolver problemáticas como impactos na agricultura, na economia e questões relacionadas à insegurança alimentar. Além disso, é importante destacar a necessidade do uso de outras culturas que possam ser aplicadas no período entressafras e utilizadas para diversificação do plantio, de forma que não se dependa exclusivamente de uma tecnologia única para obtenção de um produto final (Hausman, 2012; Lacerda et al., 2022; Nitsch, 1991; Valdes et al., 2016). Dessa forma, a investigação por outras fontes de matéria-prima e o desenvolvimento de novas tecnologias que possam neutralizar ou diminuir os problemas associados à produção de etanol tem se intensificado com a busca de formas de otimização. Nesse sentido, diversas matérias-primas vem sendo investigadas (Macedo et al., 2020), entre elas o sorgo.

O sorgo (*Sorghum bicolor* (L.) Moench) é um cereal selvagem adaptado às condições áridas e semiáridas, originário das regiões desérticas do continente africano. Pertence à família das gramíneas (Poaceae) assim como o milho, o trigo, o arroz e outros cereais, sendo o quinto cereal mais plantado no mundo. Em comparação ao milho, o sorgo possui propriedades muito similares em metabolismo, fisiologia e manejo, porém destacando-se por seu plantio mais resistente a intempéries e pragas, podendo apresentar um custo-benefício melhor em relação ao milho.

O cultivo do sorgo é de extrema importância nos continentes asiático e africano, onde é utilizado na alimentação humana. Nos países ocidentais, no entanto, é cultivado principalmente para alimentação animal. No Brasil, é utilizado principalmente para produção de grãos e forragem (Pereira Filho & Rodrigues, 2015), enquanto nos Estados Unidos cerca de um terço da colheita é destinada às biorrefinarias para a produção de etanol (RFA, 2024).

De acordo com a tabela de dados de produção e balanço de oferta e demanda de grãos da Companhia Nacional de Abastecimento (Conab), durante a safra 2022/23 foram produzidas 4,8 milhões de toneladas de sorgo em uma área total de 1,4 milhões de hectares. Os maiores produtores foram os estados de Goiás e Minas Gerais, responsáveis por 57% da produção total de sorgo no país (CONAB, 2023)

Existem cinco principais tipos de sorgo cultivados no Brasil, sendo esses o granífero, utilizado na produção de grãos, para alimentação animal, humana e produção de etanol; o sorgo forrageiro para produção de silagem e alimentação animal; o sorgo sacarino para produção de etanol; o sorgo vassoura, utilizado na produção de vassouras artesanais; o

sorgo biomassa, que pode ser utilizado como fonte de energia térmica em usinas (Pereira Filho & Rodrigues, 2015).

O sorgo granífero se destaca como uma matéria-prima alternativa promissora para a produção de etanol, apresentando diversas vantagens como: alto potencial de rendimento, baixa demanda de insumos e relativa tolerância à seca. Além disso, esse cereal possui alto teor de amido, como o milho, já tradicionalmente utilizado na produção de biocombustível.

Uma propriedade interessante dos grãos de sorgo reside na possibilidade da presença de compostos fenólicos/taninos em sua composição. Dentre as principais vantagens dos taninos nos grãos de sorgo, pode-se destacar a resistência contra o ataque de pássaros, insetos e fungos causadores de podridão, facilitando o plantio e reduzindo danos causados por organismos externos. Por outro lado, são compostos indesejáveis no sorgo cujo destino é a alimentação animal porque, além de conferir coloração indesejável à ração, provocam diminuição na palatabilidade e possuem efeito antinutricional, causado pela formação do complexo tanino-proteína, que provoca diminuição na digestibilidade em animais não ruminantes (Pereira Filho & Rodrigues, 2015).

No processo de produção do etanol os taninos podem, também, afetar a eficiência da fermentação, reduzindo a conversão de açúcares. No entanto, diferentes cultivares de sorgo apresentam teores variáveis de fenólicos/taninos. Essa variabilidade abre portas para a seleção de cultivares com menor teor desses compostos, otimizando a produção de etanol e superando os desafios impostos pela composição dos grãos (Taylor et al., 2006). Dessa forma, a propriedade de interação entre taninos e proteínas pode apresentar um desafio na produção de etanol devido ao uso de enzimas para a hidrólise do amido a açúcares simples fermentáveis (Monteiro et al., 2005; Schaffert & Rodrigues, 2014).

Por isso, mais estudos se fazem necessários para compreender a ação dos compostos fenólicos/taninos nos processos que objetivam a produção de etanol. Para que isso seja possível, o primeiro passo é a determinação dos compostos fenólicos presentes no sorgo. Os procedimentos laboratoriais frequentemente utilizados para tais determinações são métodos trabalhosos e demorados, que requerem padrões, preparos de amostra complicados e não fornecem informações adicionais sobre as amostras (Palacios et al., 2021). Portanto, novas abordagens analíticas são bem-vindas, para acelerar o processo, obter informações mais completas e precisas sobre as amostras e abrir caminho para novos avanços científicos e tecnológicos.

Nesse sentido, a metabolômica tem sido utilizada para examinar de forma abrangente misturas complexas contendo grande número de metabólitos. Aliada aos métodos de análise multivariada, como a Análise de Componentes Principais (PCA), pode ser usada como ferramenta de comparação em grandes conjuntos de dados para identificar e agrupar grupos que sejam semelhantes ou diferentes, consideradas as devidas condições, além de ser possível investigar o porquê de determinadas diferenças, caso estas sejam observadas.

## 2. OBJETIVOS

O objetivo principal desta pesquisa foi (1) avaliar o uso da metabolômica por RMN associada a métodos quimiométricos na diferenciação do perfil metabólico de quatro genótipos de sorgo granífero fornecidos pela Embrapa Milho e Sorgo, (2) tentar associar esses perfis metabólicos com melhor eficiência na produção de etanol e (3) avaliar a eficiência na produção de etanol em amostras *in natura* e após tratamento desengordurante.

### 3. REFERENCIAL TEÓRICO

#### 3.1 Metabolômica

As ciências ômicas são um campo científico que visam o estudo e caracterização estrutural de famílias de moléculas em organismos, buscando o entendimento do seu funcionamento celular e alterações biológicas. Esses estudos possuem objetivos variados, a depender do campo de pesquisa, sendo que, em todos os casos, a análise de diferentes famílias de compostos é possível dada a recente revolução instrumental (Jaumot et al., 2018).

Pensando-se nas ciências ômicas, fazem parte deste conjunto a genômica, que estuda o DNA e a informação genética dentro da célula, a transcriptômica, que estuda o RNA e as diferenças na expressão do RNAm, a proteômica, que estuda alterações das proteínas e, por fim, a metabolômica, que estuda os substratos e produtos do metabolismo, os quais são influenciados tanto pelos fatores genéticos quanto ambientais (Canuto et al., 2018; Guitton et al., 2017; Pilon et al., 2021). Considera-se metabólitos os produtos intermediários produzidos durante o metabolismo, oriundos das transformações químicas em uma amostra biológica. É chamado de metaboloma o conjunto que abrange todos os metabólitos de baixa massa molecular (até 1500 Da), sendo esses presentes em um sistema biológico ou alterados por este (Canuto et al., 2018). A metabolômica se apresenta, então, como uma ferramenta utilizada no estudo de sistemas biológicos que objetiva identificar e quantificar metabólitos presentes em um organismo ou em um dado componente deste, como resposta a fatores ambientais, interações genéticas ou de desenvolvimento (Lindon & Nicholson, 2008; Maraschin et al., 2017; Marchev et al., 2021; Patti et al., 2012).

Análises metabolômicas vêm sendo aplicada em distintas áreas de atuação, com estudos em análises clínicas, alimentos e nutrição, esportes, ambiental, toxicologia forense, análise de organismos patológicos (parasitas, bactérias, fungos), farmacologia (Canuto et al., 2018) e, de acordo com a necessidade de cada estudo, a abordagem metabolômica pode se dividir em análises classificadas como direcionadas (*target*) ou não direcionadas (*untarget*).

Análises direcionadas focam em um grupo específico onde se pretende identificar e quantificar tantos metabólitos quanto possível dentro deste espaço amostral previamente definido, sendo esta análise importante para avaliar a possível influência de um grupo de compostos de interesse na amostra, sob determinadas condições. O estudo metabolômico

seletivo requer níveis mais altos de purificação e a extração dos metabólitos almejados (Cevallos-Cevallos et al., 2009; Marchev et al., 2021). Análises não direcionadas, por outro lado, focam na detecção de tantos grupos metabólicos quanto possíveis e na comparação, em um grupo amostral, sem viés, com o objetivo de se obter perfis metabólicos e padrões/impressões digitais, chamados *fingerprints*, sem necessariamente identificar ou quantificar um composto específico (Cevallos-Cevallos et al., 2009; Dumont et al., 2020; Patti et al., 2012).

De acordo com Cevallos-Cevallos et al. (2009), tendo como base o objetivo específico da análise e o tratamento dos dados, a maioria dos estudos metabolômicos, além de direcionados e não direcionados, também podem estar inseridos em um segundo grupo onde podem ser classificados como discriminativos, informativos ou preditivos.

Análises metabolômicas discriminativas objetivam diferenças entre populações amostrais sem a necessidade de criar modelos estatísticos ou elucidar as diferenças. Normalmente faz uso de técnicas de análise multivariada de dados que ao invés de examinar cada variável isoladamente, considera a interdependência entre as variáveis, permitindo uma compreensão mais profunda e completa dos padrões e relacionamentos nos dados. Elas são aplicadas de forma a maximizar a classificação a partir do alto volume de dados, sendo a análise de componentes principais (PCA – *Principal Component Analysis*) e de agrupamento hierárquico de clusters (HCA - *hierarchical cluster analysis*) as principais ferramentas. As classificadas como informativas são focadas na identificação e quantificação de metabólitos em análises conduzidas ambas de forma direcionada ou não direcionada objetivando-se extrair informações úteis de conjuntos de dados complexos e multidimensionais, ajudando na tomada de decisões e na descoberta de conhecimentos significativos sobre a amostra.

A análise multivariada permite a consideração simultânea de múltiplas variáveis e a identificação de padrões complexos, mas pode exigir grandes conjuntos de dados, ser difícil de interpretar devido a pressupostos rigorosos e custos computacionais elevados e vem sendo aplicada na atualização de bancos de dados de metabólitos. Também se mostrou uma ferramenta útil na descoberta de biomarcadores, novos compostos bioativos e estudos de funcionalidade de metabólitos. Além destas, algumas análises metabolômicas podem ser do tipo preditivas. Nestas, são criados modelos estatísticos tendo como base o perfil e a abundância de metabólitos de forma a prever uma variável que é de difícil quantificação por outros meios, geralmente

produzidos pela ferramenta de regressão de mínimos quadrados parciais (PLS - *Partial Least Squares Regression*) (Cevallos-Cevallos et al., 2009).

Até o presente, a utilização de apenas técnicas analíticas não possibilita a obtenção do quadro metabolômico completo de uma dada amostra, sendo necessária a adoção de um conjunto interdisciplinar de abordagens que integra a biologia, a química analítica, a química orgânica, a quimiometria e a informática (Maraschin et al., 2017).

### **3.2 Técnicas analíticas**

Muitas técnicas analíticas que hoje existem podem ser aplicadas no estudo e avaliação de metabólitos, tanto qualitativa e quantitativamente. Podem ser consideradas as mais diversas técnicas, tais como ultravioleta-visível (UV-Vis), ionização por chama, espectroscopia no infravermelho (IV), espectrometria de massas (EM), ressonância magnética nuclear (RMN) e, também, técnicas cromatográficas (Patti et al., 2012).

Para análises direcionadas, frequentemente são empregadas técnicas que oferecem vantagens quanto à especificidade e reprodutibilidade quantitativa, como a cromatografia líquida (CL), cromatografia gasosa (CG), normalmente acoplada a alguma técnica de identificação como a espectrometria de massas (EM), ou a ressonância magnética nuclear quantitativa (RMN-q). Na abordagem não direcionada, são utilizadas técnicas como ressonância magnética nuclear de hidrogênio (RMN de  $^1\text{H}$ ) e carbono (RMN de  $^{13}\text{C}$ ) em abordagens uni e bidimensionais, espectroscopia vibracional de infravermelho médio ou próximo e espectrometria de massas (Maraschin et al., 2017; Patti et al., 2012).

Cada uma das técnicas analíticas listadas possui pontos fortes e fracos, relacionados tanto a variedade nas propriedades químicas e físicas dos metabólitos, como polaridade, volatilidade, estabilidade, variações na massa molar; e quanto à sua concentração na amostra. As técnicas mais comumente empregadas nos estudos metabolômicos são a espectrometria de massas e a ressonância magnética nuclear (Machado, 2014).

#### **3.2.1 Ressonância Magnética Nuclear**

A ressonância magnética nuclear é uma importante ferramenta na pesquisa metabolômica. É uma técnica analítica tão poderosa quanto as demais, além de apresentar entre as diversas vantagens a possibilidade de identificação estrutural dos metabólitos de interesse e

ser uma boa abordagem para detectar simultaneamente metabólitos primários (carboidratos, aminoácidos e lipídeos) e metabólitos secundários (compostos fenólicos, flavonoides, alcaloides, terpenos) entre outros. Mesmo sendo uma técnica menos sensível que a espectrometria de massas, estudos metabolômicos baseados em RMN se destacam por requererem preparação mínima da amostra, sendo que os espectros podem ser obtidos a partir dos extratos brutos, não exigindo etapa de derivatização; não são destrutivos, permitindo que a amostra seja posteriormente analisada por outras técnicas, eliminam a necessidade de limpeza do instrumento e evitam o risco de contaminação. As determinações por RMN são altamente reprodutíveis e a estabilidade da resposta instrumental pode durar anos se as amostras forem armazenadas de forma adequada. Além disso, o RMN de  $^1\text{H}$  pode fornecer informações quantitativas sobre metabólitos e permite a sua identificação em uma mistura por comparação dos dados com referências ou pela elucidação de uma estrutura por técnicas bidimensionais (Kim et al., 2010; Moco, 2022; Pimenta et al., 2013) Os sinais em um espectro de RMN, estabelecendo parâmetros como T1, T2, NOE, entre outros, são proporcionais à concentração de cada metabólito presente em uma amostra, de forma que ao se almejar a quantificação absoluta destes não se faz necessária a utilização de padrões separados para cada metabólito, apenas um padrão comum a todo o espectro (Last et al., 2007).

### 3.2.1.1 RMN de $^1\text{H}$

A ressonância magnética nuclear de hidrogênio é a técnica mais utilizada no âmbito de RMN. O isótopo de  $^1\text{H}$  apresenta alta abundância natural (99,98%) e, na maioria dos casos, a técnica é suficiente para gerar dados de uma amostra dentro de um tempo relativamente curto. Para os dados metabolômicos, a RNM de  $^1\text{H}$  resulta em um espectro contendo várias centenas de sinais, que reflete a quantidade de todos os metabólitos no extrato, apresentando alta reprodutibilidade e fácil quantificação, caso necessário (Kim et al., 2010; Last et al., 2007). A obtenção do espectro de RMN independe das características químicas dos metabólitos, como polaridade, acidez e basicidade. Além disso, desvantagem previamente reportada com a baixa sensibilidade pode ser contornada aumentando-se a intensidade do campo magnético, com o uso de criosondas e a miniaturização da amostragem permitindo que a sensibilidade alcance a ordem de microgramas (Pilon et al., 2020).

Em geral, espectros de RMN de  $^1\text{H}$  apresentam muitos sinais sobrepostos, o que dificulta a identificação dos sinais e a integração precisa dos picos. Essas dificuldades podem ser amenizadas fazendo-se o uso de técnicas de ressonância magnética nuclear bidimensionais,

que facilitam a determinação de constantes de acoplamento e multiplicidade dos sinais e tornam possíveis correlações  $^1\text{H}$ - $^1\text{H}$  e  $^1\text{H}$ - $^{13}\text{C}$  (Kim et al., 2010).

### 3.2.1.2 RMN bidimensional

Além do espectro unidimensional de hidrogênio existem sequências que auxiliam na identificação dos metabólitos e determinação de suas conectividades, tais como: *J*-resolvido, de correlação homonuclear (COSY – *correlation spectroscopy*), de correlação total (TOCSY – *total correlation spectroscopy*) e as de Coerência Heteronuclear de Simples Quantum (HSQC – *heteronuclear single quantum correlation*) e Correlação Heteronuclear via Múltiplas Ligações (HMBC – *heteronuclear multiple bond correlation*). Essas técnicas informam de maneira detalhada a disposição e correlação dos sistemas de *spin*, sendo atualmente utilizadas para fins de elucidação estrutural em estudos metabolômicos.

#### 3.2.1.2.1 *J*-resolvido

O experimento *J*-resolvido permite a identificação em duas dimensões dos deslocamentos químicos ( $\delta$ , ppm, eixo  $F_2$ ), as constantes de acoplamento escalares entre os núcleos e as multiplicidades ( $J$ , Hz, eixo  $F_1$ ), de modo que se faz possível examinar os parâmetros individualmente sem a interferência e as complicações do outro, como acontece em um espectro de RMN de  $^1\text{H}$ . É um experimento útil porque resolve problemas como sobreposição de sinais de multipletos vizinhos e sinais com deslocamentos químicos semelhantes (Ludwig & Viant, 2010).

#### 3.2.1.2.2 COSY

A técnica de RMN de correlação homonuclear, COSY, é uma técnica bidimensional na ressonância magnética nuclear que estabelece correlações entre hidrogênios que estejam acoplados entre si por  $^{2-3}J_{\text{H,H}}$  (Figura 1), ou seja, com acoplamentos geminais (mesmo carbono) e vicinais (carbono adjacente, permitindo discernir a multiplicidade dos sinais observados no espectro de hidrogênio RMN de  $^1\text{H}$  (Jesus, 2018; Kaiser, 2000).

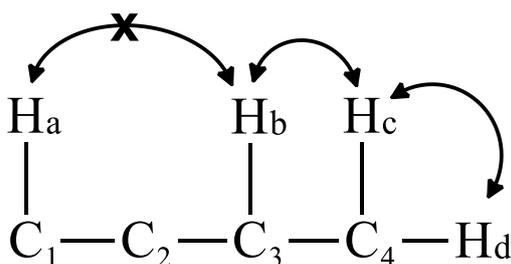


Figura 1. Acoplamentos possíveis que ocorrem no COSY: geminais (H<sub>c</sub>-H<sub>d</sub> no carbono C<sub>4</sub>) e vicinais (H<sub>b</sub>-H<sub>c</sub> nos carbonos C<sub>3</sub> e C<sub>4</sub>). A correlação resultante do acoplamento H<sub>a</sub>-H<sub>b</sub> dos carbonos C<sub>1</sub> e C<sub>3</sub> não ocorre.

### 3.2.1.2.3 HSQC e HMBC

As técnicas de ressonância magnética bidimensionais de correlação heteronuclear são utilizadas para assinalar átomos de hidrogênio correlacionados a núcleos de <sup>13</sup>C, <sup>15</sup>N e <sup>31</sup>P. São espectros editados com eixo horizontal de dimensão F<sub>2</sub> correspondendo aos deslocamentos δ<sub>H</sub> e eixo vertical de dimensão F<sub>1</sub> que correspondem aos deslocamentos para o núcleo de interesse (em geral, δ<sub>C</sub>). Acoplamentos entre átomos de hidrogênio e carbono podem ser identificados pelas sequências de pulso envolvendo transições múltiplo-quânta. Para isso, têm-se o experimento de correlação quântica única heteronuclear (HSQC - *Heteronuclear MultipleBond Coherence*) e o experimento de correlação heteronuclear de múltiplas ligações (HMBC - *Heteronuclear MultipleBond Coherence*) (Kaiser, 2000).

No HSQC são exibidas as correlações <sup>1</sup>J<sub>C-H</sub>, ou seja, determina quais núcleos de hidrogênio estão diretamente ligados a núcleos de carbonos específicos (Figura 2a). No HMBC são registradas correlações de longa distância, principalmente correlações <sup>2</sup>J<sub>C-H</sub> e <sup>3</sup>J<sub>C-H</sub>, ou seja, determina quais carbonos estão vizinhos ao hidrogênio em questão, separados por uma distância de duas, três ou mais ligações (Figura 2b). O experimento de correlação por múltiplas ligações (HMBC) também pode registrar a correlação entre hidrogênios que estejam vizinhos a carbonos não hidrogenados (carbonilas, amidas, ésteres) (Pavia et al., 2010).

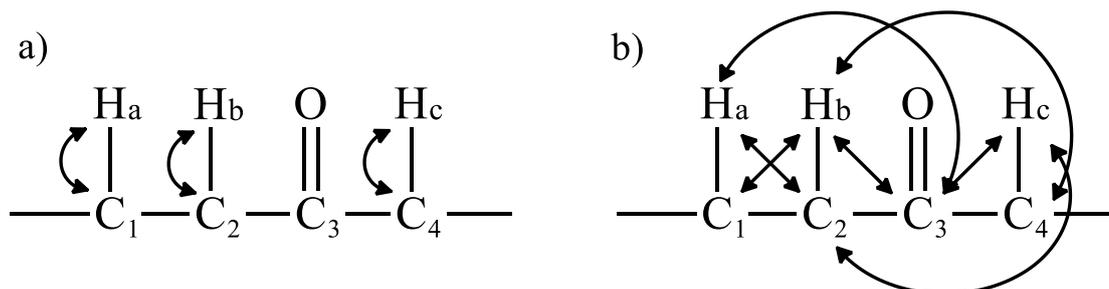


Figura 2. (a) Acoplamento direto observado pelo HSQC; e (b) possibilidades de acoplamento à longa distância observado nos experimentos de HMBC.

### 3.3 Análise multivariada de dados

Em análises metabolômicas, cada estudo é caracterizado por um grande volume de dados de características multivariadas. Em estudo multivariado consiste na análise simultânea do comportamento dos indivíduos e de suas características, considerando a combinação de múltiplas informações. Visando lidar com enormes conjuntos de dados de centenas de amostras, algumas ferramentas e métodos precisam ser aplicados para a realização do tratamento dos dados e a interpretação destes. Nesse sentido, ferramentas quimiométricas têm sido utilizadas, sendo os métodos de análise multivariada os mais indicados para explorar e extrair informações a partir de dados provenientes de estudos metabolômicos (Kim et al., 2010; Maraschin et al., 2017). A utilização da análise multivariada tem como um de seus objetivos reduzir o número de variáveis, agrupando-as baseando-se em correlações matemáticas e sistemáticas, com o objetivo de reduzir a redundância da matriz de dados e organizá-los em uma estrutura que possa facilitar a visualização de todo o conjunto (Panero et al., 2009).

Entre os métodos de análise multivariada de dados, abordagens mais populares incluem métodos não supervisionados, que modelam dados sem o uso de rótulos ou respostas pré-definidas, como a análise de componentes principais (PCA) e agrupamento hierárquico de clusters (HCA), que são aplicados rotineiramente para extrair informações de grandes conjuntos de dados e obter reconhecimento de padrões nos perfis metabólicos (Kim et al., 2010; Sumner et al., 2003).

Métodos lineares baseados em projeções como o PCA são populares em metabolômica pela simplicidade da interpretação e assim são usados para o entendimento de eventos biológicos. No entanto, métodos não lineares como redes neurais, máquina de vetores de

suporte, *random forest* são menos utilizados em metabolômica, sendo restritos a previsão de variáveis dentro de um contexto de classificação.

### 3.3.1 Análise de Componentes Principais (PCA)

A PCA é uma ferramenta usada para reduzir a dimensionalidade dos dados ao mesmo tempo em que preserva a maior quantidade possível de variância presente no conjunto de dados original com a menor perda possível da informação. É um método de agrupamento que não requer conhecimento do conjunto e atua reduzindo a dimensionalidade dos dados multivariados, preservando a maior parte da variação dentro deles. A técnica permite descrever a variância em um conjunto multivariado de dados metabolômicos em *plots* gráficos, visualizações e agrupamentos com base em combinações lineares, conhecidas como componentes principais (PC – *Principal Component*). As componentes principais são combinações lineares das variáveis originais e são independentes entre si, permitindo extrair as informações com um número reduzido de variáveis. Como técnica de agrupamento, é mais comum a PCA ser utilizada para identificar como uma amostra se difere de outra e quais variáveis mais contribuem para essa diferença (Choi et al., 2004; Wishart, 2008).

Cada componente principal é caracterizada por duas informações principais: as pontuações (*scores*) que, no caso da ressonância magnética nuclear, referem-se aos perfis e particularidades das amostras, e as cargas ou pesos (*loadings*), que se referem aos dados espectrais de deslocamento químico (Brereton, 2003). A PCA gera um plano modelo que fornece o gráfico de pontuações (*scores*) onde é visualizada a relação entre as observações ou amostras, ou seja, observa-se a existência de quaisquer agrupamentos, tendências ou valores discrepantes no conjunto de dados analisado. Os scores podem ser apresentados graficamente, em um gráfico de pesos (*loadings*) que descreve a influência das variáveis no modelo e a relação entre elas (Trygg et al., 2007).

A PCA é a técnica quimiométrica multivariada de utilização mais difundida e, devido à importância da aplicação de análises multivariadas em química, é considerada por muitos como a técnica que mais impactou a visão do químico acerca da análise de dados (Brereton, 2003).

Espectros obtidos por ressonância magnética nuclear de hidrogênio (RMN de  $^1\text{H}$ ), principalmente para amostras vegetais, oferecem um número extenso de sinais, o que dificulta a obtenção de informações a partir apenas da inspeção visual destes. Apesar disso, os espectros ainda são ferramentas importantes e ricas em informações sobre a constituição da amostra.

Através de métodos quimiométricos, é possível fazer o processo de correlação entre as informações contidas no espectro e as propriedades da amostra que o gerou (Wishart, 2008).

Em uma análise por PCA é possível observar distinções entre grupos ou a ocorrência de discriminação de variáveis. Essas informações, quando obtidas a partir dos espectros de RMN, podem indicar quais substâncias caracterizam um determinado grupo de amostras, seja por variações de concentração ou pela presença/ausência de um composto, o que pode inferir na identificação de biomarcadores (Lindon & Nicholson, 2008).

### **3.3.2 Agrupamento Hierárquico de Clusters (HCA)**

A análise de agrupamento hierárquico de clusters trata-se de um método não supervisionado de reconhecimento de padrões que é adequado para descobrir padrões naturais de distribuição das amostras baseando-se no perfil multivariado. A análise HCA é realizada a partir da redução de variáveis e dados redundantes pelo processo de análise PCA e tem como resultado uma árvore de agrupamentos que apresenta divisões dos dados amostrais em grupos (*clusters*) e tem como objetivo exibir os dados em um espaço bidimensional enfatizando agrupamentos e padrões naturais. Na raiz da árvore estão todos os dados agrupados em um único grupo, que vai se dividindo de forma binária, ou seja, dois grupos são criados a partir da divisão de um anterior. O processo relaciona amostras de forma que as semelhantes são agrupadas entre si, resultando em um dendrograma que separa amostras ou variáveis em função de suas similaridades. Esse dendrograma varia da escala zero, representando amostras sem similaridades, até um, que representa as amostras similares (Ferreira, 2015; Maraschin et al., 2017; Panero et al., 2009).

## 4. METODOLOGIA

### 4.1 Amostras utilizadas

Quatro amostras de *Sorghum bicolor* L. Moench foram cedidas pela Embrapa Milho e Sorgo, localizada em Sete Lagoas/MG (Figura 3), para o desenvolvimento do projeto.



Figura 3. Amostras de sorgo granífero.

As amostras são de sorgo granífero, sendo híbridos (BRS 305 e BR501) e linhagens (SC084 e CMSXS 180), com teores variados de compostos fenólicos (Tabela 1). Os grãos foram previamente moídos na Embrapa em uma granulometria fina de 0,5 mm.

Tabela 1. Descrição das amostras de *Sorghum bicolor* L. Moench granífero cedidas pela Embrapa Milho e Sorgo para desenvolvimento do projeto de pesquisa

| Amostra  | Variedade | Safra | Coloração do pericarpo | Taninos <sup>1</sup> |
|----------|-----------|-------|------------------------|----------------------|
| SC084    | Linhagem  | 2018  | Vermelho               | Contém               |
| BRS305   | Híbrido   | 2020  | Vermelho               | Contém               |
| CMSXS180 | Linhagem  | 2018  | Branco                 | Não contém           |
| BR501    | Híbrido   | 2007  | Branco                 | Não contém           |

<sup>1</sup> Informação prévia sobre as cultivares enviada pela Embrapa junto com as amostras. A informação “contém taninos” e “não contém taninos” foi fornecida, sem quantificações.

## 4.2 Tratamento das amostras: desengorduramento dos grãos

Para a obtenção das amostras desengorduradas utilizaram-se 50,0 g das amostras *in natura*. Foram adicionados 100,0 mL de hexano às amostras, seguido de aquecimento a 40 °C sob agitação magnética por 30 minutos, extração em banho de ultrassom durante 20 minutos e filtração à vácuo. O procedimento de extração foi repetido 4 vezes para cada amostra. O filtrado foi levado para o evaporador rotativo sob vácuo (600 mbar) e aquecimento (50 °C) para recuperação do solvente e obtenção do extrato hexânico. As amostras de sorgo desengorduradas foram mantidas na capela sob exaustão até completa evaporação do solvente, quando foram pesadas, acondicionadas em frascos adequados e mantidas sob refrigeração.

## 4.3 Caracterização química das amostras de sorgo

### 4.3.1 Determinação do teor de compostos fenólicos totais

A determinação do teor de compostos fenólicos totais das amostras de sorgo foi realizada a partir de duas metodologias baseadas na utilização do reagente Folin-Ciocalteu. O teste de Folin-Ciocalteu é largamente usado para quantificar fenóis e polifenóis em diferentes matrizes. Apesar de ser um método sensível, não é específico para um único grupo de compostos fenólicos, sendo utilizado para quantificar a concentração total de compostos fenólicos em extratos. O método parte de uma reação de oxirredução onde as hidroxilas fenólicas são oxidadas pelo ácido fosfomolibdico-fosfotúngstico do reagente de Folin-Ciocalteu, originando um complexo de coloração azul vívido que pode ser medido a 765nm (Angelo & Jorge, 2007; Pérez et al., 2023; Schofield et al., 2001).

#### 4.3.1.1 Método 1

Em um tubo Falcon de 50 mL foram pesadas  $0,2500 \pm 0,0025$  g de amostra. Ao tubo, foram adicionados 25,0 mL de solução ácida de metanol (HCl 1% v/v) e o mesmo foi mantido sob agitação em *shaker* a 200 r.p.m. durante 2 horas. Em seguida, os tubos foram centrifugados a 3000 r.p.m. por 10 minutos. Do sobrenadante obtido, 100 µL foram transferidos para um tubo Falcon de 15 mL, ao que foram adicionados 1,1 mL de água deionizada, 0,4 mL de solução reagente Folin-Ciocalteu 0,02 N, 0,9 mL de solução de etanolamina 0,5 mol L<sup>-1</sup>. O tubo foi agitado em vórtex durante 10 segundos e mantido em repouso durante 20 minutos para a etapa

de reação. Foi preparado um branco para controle negativo adicionando-se todos os reagentes listados a 100  $\mu\text{L}$  de metanol P.A.

As leituras foram realizadas no espectrofotômetro UV-Visível 190-1100 nm da Kasvi no comprimento de onda de 600 nm. Construiu-se uma curva padrão de absorbância em função da concentração de ácido gálico nas concentrações de 5, 25, 50, 100, 150, 200 e 250  $\mu\text{g mL}^{-1}$ . O resultado final obtido foi expresso em equivalentes de ácido gálico ( $\text{mg GAE g}^{-1}$  amostra).

#### 4.3.1.2 Método 2

Em um Eppendorf de 2,0 mL foram pesados 0,100 g das amostras de sorgo. Ao Eppendorf, adicionou-se 1,0 mL de metanol grau HPLC e 1,0 mL de água deionizada. O Eppendorf foi agitado em vórtex por 1 minuto, levado ao banho de ultrassom por 20 minutos e, em seguida, centrifugado a 16000 r.p.m. por 15 minutos. Ao final, o sobrenadante foi utilizado no teste.

Para a determinação de compostos fenólicos totais, foram preparadas soluções em diferentes concentrações utilizando-se alíquotas de 5,0; 10,0; 20,0; 50,0; 100,0 e 200,0  $\mu\text{L}$  do extrato previamente obtido diluídas para 1,0 mL com água destilada. O mesmo procedimento foi realizado com ácido gálico como controle positivo. Em uma microplaca de 96 poços foram adicionados 80  $\mu\text{L}$  das soluções preparadas a partir do extrato das amostras, da solução padrão de ácido gálico 1,0  $\text{mg mL}^{-1}$  e uma solução metanol/água (1:4) como branco para controle negativo. Todas as amostras foram testadas em triplicata. Para as amostras informadas com ausência de compostos fenólicos (BR501 e CMSXS180) também foi adicionado o extrato sem diluição a um dos poços da placa, sem triplicata, para teste.

A determinação do teor de compostos fenólicos foi realizada por método espectrofotométrico em espectrofotômetro UV-Vis Multiskan GO da Thermo Scientific no comprimento de onda de 760 nm. A absorbância inicial dos extratos foi determinada e, em seguida, adicionaram-se 80  $\mu\text{L}$  de solução de Folin-Ciocalteu 1:3 (reagente/água) em cada poço e a mistura foi mantida em repouso por 5 minutos no interior do equipamento, para a reação ser realizada sem a incidência de luz. A absorbância foi novamente determinada. Adicionaram-se 120  $\mu\text{L}$  de solução de carbonato de sódio 35% e, utilizando-se a função “agitar” do equipamento, a placa foi mantida sob agitação constante em velocidade média durante 20 segundos. Por fim, foram adicionados 120  $\mu\text{L}$  de água destilada e a absorbância final foi determinada.

Construiu-se um gráfico de absorvância em função da concentração de ácido gálico nas concentrações de 25, 50, 100, 150 e 200  $\mu\text{g mL}^{-1}$  e o resultado final obtido foi expresso em equivalentes de ácido gálico ( $\text{mg GAE g}^{-1}$  amostra).

#### 4.3.2 Determinação do teor de taninos pelo método Vanilina-HCl

A análise de determinação de taninos pelo método Vanilina-HCl foi realizada seguindo-se adaptações na metodologia de Price, Van Scoyoc e Butler (1978). A determinação visa a quantificação de proantocianidinas, que são taninos condensados usualmente encontrados no pericarpo do sorgo. O método da vanilina depende da reação da vanilina com os taninos condensados do extrato, originando a formação de complexos de cor vermelha (Figura 4) (Rao et al., 2018; Schofield et al., 2001).

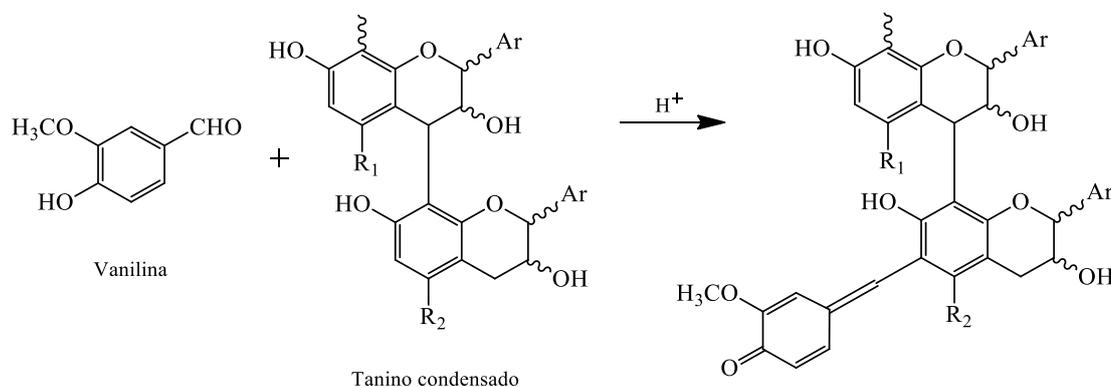


Figura 4. Equação representativa da reação da vanilina com um tanino condensado.

Em um tubo Falcon de 50 mL foram pesadas, em triplicata, 0,3000 a 0,3030 g das amostras com alto teor de taninos (SC084 e BRS305) e 0,1500 a 0,1515 g das amostras com baixo teor de taninos (CMSXS180 e BR501). Aos tubos foram adicionados 8,0 mL de solução de HCl 1% (v/v) em metanol seguido de agitação em vórtex por 10 segundos. As amostras foram levadas para o agitador *shaker* a 200 r.p.m. por 20 minutos e, após, para centrifuga a 3000 r.p.m. por 15 minutos. De cada amostra foi retirada uma alíquota de 1,0 mL, em duplicata, que foram transferidas para tubos Falcon de 15 mL previamente identificados.

À primeira alíquota (amostras) foram adicionados 5,0 mL de solução de vanilina (1:1 Vanilina 1% : HCl 8% em metanol) e os tubos foram agitados em vórtex por 10 segundos. À segunda alíquota (branco da amostra) adicionaram-se 5,0 mL de solução de HCl 4% em metanol e agitou-se em vórtex por 10 segundos. Os tubos foram mantidos em repouso durante 20 minutos e, em seguida, foi realizada a leitura a 500 nm no espectrofotômetro UV-Vis 190-1100 nm da Kasvi. Metanol PA foi utilizado para correção das absorvâncias (branco).

Construiu-se um gráfico de absorvância em função da concentração de catequina (CE) nas concentrações de 0,25; 0,50; 0,75; 1,00 e 1,50 mg mL<sup>-1</sup> e o resultado final obtido foi expresso em equivalentes de catequina (mg CE g<sup>-1</sup> amostra).

#### 4.3.3 Determinação do teor de amido total

Durante a produção do etanol, é feita a hidrólise enzimática do amido para a obtenção de açúcares fermentáveis. Dessa forma, o teor de amido total é um dado importante para o cálculo do rendimento teórico do processo de fermentação e, conseqüentemente, na determinação da eficiência da conversão do processo de produção de etanol.

A determinação do teor de amido total foi feita no Laboratório de Composição Centesimal (LCC) da Embrapa Milho e Sorgo utilizando o Kit Megazyme para análise de amido total. O kit é composto por enzimas  $\alpha$ -amilase (3000 U mL<sup>-1</sup>) e glucoamilase (3300 U mL<sup>-1</sup>), enzima reagente GOPOD (enzima glicose oxidase e peroxidase), reagente tampão GOPOD (ácido *p*-hidroxibenzoico e azida de sódio), glicose e padrão de amido.

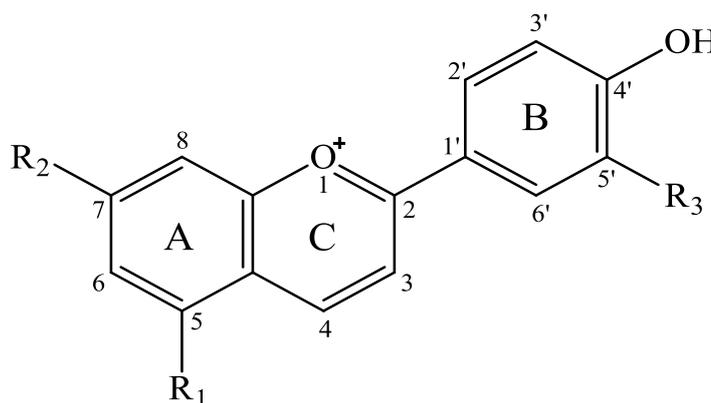
Em tubo de ensaio foram pesados aproximadamente 0,1000 g da amostra de sorgo. Ao tubo, adicionaram-se 200  $\mu$ L de solução aquosa de etanol (80% v/v) e agitou-se em vórtex. Em seguida, adicionaram-se 3,0 mL de solução  $\alpha$ -amilase (1,0 mL de  $\alpha$ -amilase 3000 U mL<sup>-1</sup> em 30,0 mL de solução de acetato de sódio 100 mmol L<sup>-1</sup>). Os tubos de ensaio foram incubados em banho-maria fervente por 6 minutos sob agitação vigorosa a cada 2 minutos, seguido de resfriamento a temperatura ambiente por 10 minutos. Após esse período, os tubos foram novamente levados ao banho-maria a 50 °C por 2 minutos. A cada tubo foram adicionados 100  $\mu$ L de enzima glucoamilase 3300 U mL<sup>-1</sup>, seguido de agitação em vórtex e incubação a 50 °C em banho-maria por 30 minutos. Durante o período de incubação, os tubos foram agitados a cada 10 minutos em vórtex. O conteúdo de cada tubo foi vertido para um balão volumétrico de 100,0 mL e o volume foi ajustado com água deionizada. Do conteúdo do balão volumétrico, 30,0 mL foram transferidos para um tubo Falcon de 50 mL, que foi centrifugado a 4000 r.p.m. durante 10 minutos. Duas alíquotas de 100  $\mu$ L do sobrenadante foram transferidas para tubos de ensaio, intitulados A e B. A um dos tubos foram adicionados 3,0 mL de solução GOPOD (enzima glicose oxidase e peroxidase em tampão ácido *p*-hidroxibenzoico e azida de sódio) e ambos foram incubados em banho-maria a 50 °C por 20 minutos.

Como controle, foi preparada uma solução de glicose adicionando-se 100  $\mu$ L de solução de glicose 1 mg mL<sup>-1</sup> e 3 mL de solução GOPOD a um tubo de ensaio. Como branco para controle negativo foi preparada uma solução de 100  $\mu$ L de água deionizada e 3 mL de solução

GOPOD. O branco foi utilizado para zerar o espectrofotômetro UV-Visível 190-1100 nm da Kasvi e foram feitas as leituras do controle de glicose e das amostras no comprimento de onda de 510 nm.

#### 4.3.4 Determinação do perfil de antocianidinas por CLAE

O *Sorghum bicolor* variedade graniferum apresenta ampla gama de colorações, podendo ter grãos em tons claros/brancos até grãos com tonalidades escuras (vermelho, marrom, preto). Esta coloração é frequentemente associada à presença de polifenóis pertencentes às classes dos flavonoides e ácidos fenólicos. Dentre os flavonoides, a subclasse das antocianidinas é uma das mais importantes. Objetivando avaliar o teor de antocianidinas nas amostras de sorgo, as antocianidinas apigeninidina, 7-metoxi-apigeninidina, luteolinidina, 5-metoxi-luteolinidina e 5,7-dimetoxi-luteolinidina (Figura 5), previamente destacadas em estudos sobre as propriedades do sorgo (Awika et al., 2004), foram quantificadas por Cromatografia Líquida de Alta Eficiência (CLAE).



- $R_1 = \text{OH}, R_2 = \text{OH}, R_3 = \text{H}$ : apigeninidina  
 $R_1 = \text{OH}, R_2 = \text{OH}, R_3 = \text{OH}$ : luteolinidina  
 $R_1 = \text{OH}, R_2 = \text{OCH}_3, R_3 = \text{H}$ : 7-metoxi-apigeninidina  
 $R_1 = \text{OCH}_3, R_2 = \text{OH}, R_3 = \text{OH}$ : 5-metoxi-luteolinidina

Figura 5. Estrutura das antocianidinas identificadas nos grãos de sorgo

Para a determinação do perfil de antocianidinas, foi feita a extração de  $(1,000 \pm 0,001)$  g de amostra em 20,0 mL de solução HCl 1% em metanol (v/v). Para a extração, a solução foi agitada durante 2 horas a 200 r.p.m. em um *shaker*, seguido de centrifugação a 3000 r.p.m. por 10 minutos. O extrato centrifugado foi filtrado em um filtro de seringa hidrofílico de 0,22  $\mu\text{m}$ , dos quais 1,5 mL foram transferidos para um *vial* e utilizados para injeção no equipamento. Além das amostras extraídas, também foram analisados cinco padrões: luteolinidina,

apigeninidina, 7-metoxi-apigeninidina, 5,7-dimetoxi-luteolinidina e 5,7-dimetoxi-apigeninidina.

As análises foram feitas no laboratório da Embrapa Milho e Sorgo por cromatografia líquida de alta eficiência (CLAE) em um cromatógrafo da marca Waters modelo e2695 utilizando coluna C18 de fase reversa sob a temperatura de 35 °C e detector de arranjo de diodos (DAD) em um comprimento de onda de 480 nm. Foram construídas curvas de absorvância em função da concentração de padrões de antocianidinas e o resultado final para as amostras foram expressos em  $\mu\text{g g}^{-1}$ .

#### **4.3.5 Determinação da composição centesimal**

A determinação dos teores de carboidratos, celulose, hemicelulose, extrato etéreo, fibras, lignina, matéria mineral, proteínas e umidade foram realizadas no Laboratório de Composição Centesimal da Embrapa Milho e Sorgo após colheita e processamento dos grãos de sorgo segundo metodologias oficiais descritas pela Associação de Químicos Analíticos Oficiais (Cunniff & Washington, 1995).

#### **4.4 Caracterizações espectroscópicas do sorgo**

##### **4.4.1 Determinação do perfil no ultravioleta**

Em um tubo Eppendorf foram pesadas 50,0 mg das amostras de sorgo *in natura*. Ao tubo, foram adicionados 0,75 mL de água deionizada e 0,75 mL de metanol P.A. Em seguida, o tubo Eppendorf foi levado a agitação em vórtex por 1 minuto, seguido por extração em banho de ultrassom por 20 minutos, ambos a temperatura ambiente. Por fim, a solução foi centrifugada a 16000 r.p.m. por 20 minutos para obtenção de um sobrenadante limpo. O sobrenadante obtido foi utilizado para determinação do perfil no ultravioleta.

A determinação do perfil no ultravioleta foi realizada por método espectrofotométrico em espectrofotômetro UV-Vis Multiskan GO da Thermo Scientific abrangendo os comprimentos de onda entre 200 e 600 nm.

##### **4.4.2 Determinação da atividade antioxidante**

A avaliação da atividade antioxidante das amostras de sorgo *in natura* foi realizada a partir método ABTS (2,2-azino-bis-ácido-3-etilbenzotiazolína-6-sulfônico). Esse método

baseia-se na redução do radical ABTS<sup>•+</sup>, de coloração verde escura, a ABTS, de coloração verde clara, quando ocorre a adição de um antioxidante ao meio reacional. O radical ABTS<sup>•+</sup> possui absorção máxima em 734 nm e a extensão da descoloração do meio é o indicativo da inibição do radical, que é determinada em função do padrão Trolox (ácido 6-hidroxi-2,5,7,8-tetrametilcromano-2-carboxílico), que é submetido às mesmas condições de análise das amostras (Awika et al., 2003).

Inicialmente preparou-se a solução estoque do radical ABTS, que deve ser feita no mínimo 12 horas antes da análise. Para isso, 44,0 mg do reagente ABTS foram pesadas e transferidas para um balão volumétrico de 10,0 mL, e 40,0 mg de persulfato de potássio (K<sub>2</sub>S<sub>2</sub>O<sub>8</sub>) foram pesados e vertidos em um balão volumétrico de 50,0 mL. Ambos tiveram seus volumes completados com água deionizada. A um frasco âmbar foram adicionados 10,0 mL da solução de ABTS e 10,0 mL da solução de persulfato de potássio. Para o preparo da solução de trabalho, a um novo frasco âmbar foram transferidos 5,0 mL da solução estoque do radical ABTS e 145,0 mL de solução tampão fosfato (pH 7,4). Com o auxílio de um espectrofotômetro, a absorbância da solução foi ajustada para o valor de 1,5 ( $\pm 0,05$ ) no comprimento de onda de 734 nm.

Para o preparo e leitura, foram pesados  $0,2000 \pm 0,0020$  g das amostras de sorgo em um tubo Falcon de 50 mL, em triplicata, e adicionaram-se 10 mL de solução ácida de metanol (HCl 1% v/v). O tubo foi agitado a 200 r.p.m. durante 2 horas em um *shaker*. Em seguida, foi levado a uma centrifuga a 3000 r.p.m. por 10 minutos. Do sobrenadante obtido, 100  $\mu$ L foram transferidos para um tubo Falcon de 15 mL, ao que foram adicionados 2,9 mL de solução de trabalho do radical ABTS e agitou-se em vórtex por 10 segundos. A amostra foi mantida em repouso na ausência de luz durante 30 minutos para processar a etapa de reação. Também foi preparado um branco para controle negativo, utilizando-se 100  $\mu$ L de metanol 1% HCl (v/v) e a solução de trabalho do radical ABTS.

As leituras foram realizadas em espectrofotômetro UV-Vis no comprimento de onda de 734 nm. Foi construída uma curva de absorbância em função da concentração do padrão Trolox (TE) nas concentrações de 200, 400, 600, 800 e 1000  $\mu$ mol mL<sup>-1</sup> e o resultado final obtido foi expresso em equivalentes de trolox ( $\mu$ mol TE g<sup>-1</sup> amostra) de amostra seca.

### 4.4.3 Determinação do perfil metabólico por RMN

#### 4.4.3.1 Preparo das amostras

O preparo das amostras de sorgo *in natura* e desengorduradas para a análise de Ressonância Magnética Nuclear foi realizado tendo como referência o trabalho de Kim, Choi e Verpoorte (2010). Em um tubo Eppendorf de 2 mL foram pesadas 50,0 mg das amostras. Ao tubo foram adicionados 0,75 mL de metanol- $d_4$  e 0,75 mL de tampão fosfato de potássio monobásico ( $\text{KH}_2\text{PO}_4$ ) em  $\text{D}_2\text{O}$  (pH 6,0) contendo 0,01% (m/m) de sal de sódio do ácido 3-(trimetilsilil)propionico-2,2,3,3- $d_4$  (TSP). O tampão foi preparado a partir da adição de 1,232 g de  $\text{KH}_2\text{PO}_4$  e 10,0 mg do TSP a 100 mL de água deuterada ( $\text{D}_2\text{O}$ ). A solução foi agitada até completa dissolução e o pH foi ajustado para 6,0 utilizando hidróxido de sódio deuterado ( $\text{NaOD}$ ) 1,0 mol  $\text{L}^{-1}$ . Em seguida, o tubo Eppendorf foi levado a agitação em vórtex por 1 minuto, seguido por extração em banho de ultrassom por 20 minutos, ambos a temperatura ambiente. Por fim, a solução foi centrifugada a 16000 r.p.m. por 20 minutos para obtenção de um sobrenadante limpo, do qual 800  $\mu\text{L}$  foram transferidos para um tubo de RMN previamente identificado. Todo o procedimento foi realizado em triplicata para cada uma das amostras.

#### 4.4.3.2 Obtenção dos espectros

Os experimentos de RMN foram obtidos no espectrômetro Bruker Avance Neo 600 MHz no laboratório de Ressonância Magnética de Alta Resolução (LAREMAR) do Departamento de Química/ICEx da UFMG. Todas as análises foram realizadas utilizando-se uma sonda multinuclear PA BBO 600S3 a 25 °C. Para as amostras de sorgo *in natura* e desengorduradas de todas as cultivares foram obtidos espectros de RMN de  $^1\text{H}$  e para as duas amostras linhagem (SC084 e CMSXS180), *in natura*, com maior e menor teor de compostos fenólicos foram obtidos espectros bidimensionais.

Os espectros de RMN de  $^1\text{H}$  foram adquiridos utilizando os parâmetros descritos por Kim, Choi e Verpoorte (2010) com pré-saturação do sinal de água utilizando a sequência de pulsos zgpcpr e janela espectral de 10 kHz (16,7 ppm). Para a aquisição foram utilizados 64k pontos de dados, 128 scans, tempo de aquisição (AQ) de 3,2 s e recuperação (d1) de 3,2 s. Após transformada de Fourier, as fases (zero e primeira fase) foram manualmente corrigidas e as linhas de base foram corrigidas automaticamente utilizando o programa TopSpin 4.2.0. Por fim, os espectros foram calibrados de acordo com o sinal do TSP- $d_4$  em 0,00 ppm.

Os experimentos de  $J$ -resolvido foram adquiridos com pré-saturação do sinal de água durante o tempo de relaxação ( $d_1$ ) de 2 s. O espectro foi adquirido usando uma matriz de dados de 128 x 2048 (TDF1 x TDF2) cobrindo uma janela de 120 Hz em F1 (acoplamento spin-spin) e 13,2 kHz em F2 (eixo do deslocamento químico) com 16 scans a cada incremento. Foi aplicada a transformada de Fourier, os dados transformados foram simetrizados e o espectro foi calibrado de acordo com o sinal do TSP- $d_4$  em 0,00 ppm.

Para as correlações de deslocamentos homonucleares  $^1\text{H}$ - $^1\text{H}$  no COSY, os mapas de contorno foram adquiridos utilizando uma sequência de pulsos com tempo de relaxação ( $d_1$ ) de 2s, com aquisição de dados em 1024 incrementos e 8 scans por incremento em uma janela espectral de 16,7 kHz em ambas as dimensões. Foi aplicada a transformada de Fourier, a simetria dos dados e a calibração do espectro pelo sinal do TSP- $d_4$  em 0,00 ppm.

Para as correlações de deslocamentos heteronucleares  $^1\text{H}$ - $^{13}\text{C}$  HSQC, os mapas de contorno foram adquiridos em uma matriz de 600 x 2048 (TDF1 x TDF2) pontos cobrindo uma janela espectral F1 e F2 de 37,7k Hz e 10k Hz respectivamente usando 150 scans para cada incremento com tempo de relaxação de 1,5 s. Para a amostra CMSXS180 foi obtido um HSQC seletivo utilizando uma janela reduzida. O espectro foi calibrado de acordo com o sinal do padrão interno TSP- $d_4$  em 0,00 ppm.

Para as correlações de deslocamentos heteronucleares  $^1\text{H}$ - $^{13}\text{C}$  HMBC, os mapas de contorno foram adquiridos em uma matriz de 256 x 2048 (TDF1 x TDF2) pontos, cobrindo uma janela espectral F1 e F2 de 36,2k Hz e 10k Hz, respectivamente, usando-se 150 scans para cada incremento com tempo de relaxação de 1,5 s. O espectro foi calibrado de acordo com o padrão interno TSP- $d_4$  0,00 ppm.

## **4.5 Obtenção do etanol de sorgo**

### **4.5.1 Preparo das amostras**

O etanol foi obtido pelo processo de sacarificação e fermentação simultâneas dos grãos de sorgo. Para a produção do etanol, as amostras foram previamente moídas em granulometria fina de 0,5 mm. Todos os diferentes cultivares, com diferentes teores de taninos, dos dois grupos de trabalho (*in natura* e grãos desengordurados) foram utilizados na obtenção do etanol.

#### 4.5.2 Preparo do mosto

Balões de fundo chato, previamente secos em estufa a 105 °C por 1 hora, foram identificados e tiveram as suas massas determinadas para posterior cálculo do rendimento. A cada balão foram adicionados 7,50 g de sorgo moído, em triplicata. Em seguida, foram adicionados 30,0 mL de tampão acetato 100 mmol L<sup>-1</sup> (pH 5,0).

#### 4.5.3 Liquefação

A definição de quantidade das enzimas Liquozyme® Supra 2.2X (Novozymes), responsável pela hidrólise do amido nas ligações  $\alpha$ -(1,4)-D-glicosídeo, e da enzima Spirizyme® Fuel Ultra XHS (Novozymes), uma glucoamilase que hidrolisa as ligações  $\alpha$ -1,6 da amilopectina, foi determinada tendo como referência as condições de produção de etanol a partir do milho propostas por Ramchandran e colaboradores (2015). Assim, para hidrólise do amido do sorgo, foram utilizados 6  $\mu$ L de Liquozyme® (etapa de liquefação) e 10  $\mu$ L de Spirizyme® (etapa de sacarificação) para 30 mL de mosto com 25% (m/v) de sorgo moído.

Dessa forma, após preparo do mosto, foram adicionados 6  $\mu$ L da enzima  $\alpha$ -amilase Liquozyme® (Novozyme) a cada balão. O conteúdo foi homogeneizado e o balões foram levados a banho-maria sob temperatura de 90 °C e agitação durante 3 horas, protegidos do ambiente externo com o auxílio de papel alumínio.

#### 4.5.4 Sacarificação e fermentação simultâneas

Aproximadamente 20 minutos antes do final do processo de liquefação, as leveduras foram preparadas adicionando-se 3,00 g de levedura comercial (*Saccharomyces cerevisiae*) a 15,0 mL de água. Para ativação, as leveduras foram mantidas sob agitação a 32 °C durante 30 minutos. Finalizado o processo de liquefação, os balões foram retirados do banho-maria e colocados sob a bancada para resfriarem a temperatura ambiente. Em seguida, a cada balão foram adicionados 200  $\mu$ L de solução de ureia a 50% (m/v), 1000  $\mu$ L de levedura ativada, aproximadamente 0,010 g de antibiótico tiosulfato de neomicina e 10  $\mu$ L da enzima glucoamilase Spirizyme Ultra XHS. Os balões foram fechados com rolhas de silicone acopladas a *airlocks* (Figura 6) e levados para a incubadora *shaker* pelo período de 72 horas a 32 °C e rotação de 45 r.p.m. Após 72 horas, o processo de sacarificação e fermentação simultâneas foi interrompido, as amostras fermentadas foram filtradas em cadinhos de vidro com placa porosa.



Figura 6. Balões fechados com rolhas de silicone acopladas a *airlocks* em incubadora *shaker* preparados para o processo de sacarificação e fermentação simultâneas.

O líquido obtido como resultado da fermentação foi armazenado em tubos Falcon de 15 mL e mantido em freezer para posterior determinação do teor de etanol. O resíduo retido na filtração foi recolhido para compor o resíduo da destilação (DDGS – *Dried Distillers Grains with Solubles*), que foi seco em estufa sob temperatura de 60 °C por aproximadamente 60 h. Os balões utilizados para a fermentação, com resíduos aderidos às paredes, foram levados a estufa de circulação a 60 °C por aproximadamente 60 h, seguido de resfriamento em dessecados e, por fim, pesagem para determinação da massa final. A diferença de massa dos balões antes de colocar a amostra e após filtração do DDGS caracteriza a parte da amostra não fermentada que aderiu à superfície do balão durante o processo e é importante para o cálculo do rendimento final de cada amostra (demonstração do cálculo: apêndice C, página 85).

#### 4.5.5 Determinação do teor de etanol

Os teores de etanol, glicose e glicerol foram quantificados no laboratório da Embrapa Milho e Sorgo por cromatografia no cromatógrafo líquido de alta eficiência de marca Waters modelo 22695 equipado com a coluna Aminex HPX-87H (Bio-rad). Os experimentos foram realizados utilizando como fase móvel uma solução aquosa de ácido sulfúrico 0,005 mol L<sup>-1</sup> e

fluxo de 0,5 mL min<sup>-1</sup> sob a temperatura de 50 °C e detector de índice de refração modelo Water 2414. Para a quantificação dos teores de etanol, glicose e glicerol foram construídas curvas analíticas com seis concentrações diferentes a partir de cada padrão (Sigma-Aldrich) com 99,5% de pureza. Os resultados das amostras analisadas foram expressos em mg mL<sup>-1</sup>.

#### 4.6 Análise multivariada dos dados de RMN de <sup>1</sup>H

A análise multivariada dos dados de RMN de <sup>1</sup>H foi efetuada usando módulos de análise estatística do *software* livre MetaboAnalyst 6.0 (disponível em versão web em <http://www.metaboanalyst.ca/MetaboAnalyst>). Nos últimos anos o *software* tem-se afirmado como uma boa opção para o tratamento de dados em análise metabolômica, disponibilizando uma ampla gama de operações de pré-processamento de dados e de análises uni e multivariadas. Também é uma ferramenta de fácil utilização e acesso, uma vez que é disponível a partir de um *website* comum sem a obrigatoriedade de baixar ou instalar aplicativos (Xia et al., 2009).

O pré-tratamento dos espectros de RMN de <sup>1</sup>H consistiu, inicialmente, na redução da dimensionalidade dos dados utilizando-se o processo de *binning*. No *binning*, os espectros são divididos em pequenas regiões ou intervalos representados por uma faixa específica de deslocamento químico (ppm) (chamados *bins*) e a área total de cada *bin* é calculada de forma a representar o espectro original. A abordagem mais comum consiste em dividir os espectros em janelas espectrais uniformes, em geral de 0,04 ppm. Dessa forma, os dados espectrais dos extratos das diferentes cultivares foram convertidos em uma matriz de área com o intervalo de 0,04 ppm com o uso do programa Chenomx NMR Suite 10.0. Antes das análises, as regiões  $\delta$  4,74-4,90 e  $\delta$  3,26-3,38 foram removidas para eliminar quaisquer interferências devido ao sinal do solvente (metanol) e à supressão do sinal da água.

Durante a aplicação da matriz no MetaboAnalyst, seguindo parâmetros padrão do *software* baseado em regras empíricas dos desenvolvedores, foi aplicado um filtro de 5% para identificar e remover variáveis que provavelmente não seriam úteis ao modelar os dados. A normalização foi realizada com a área total de cada espectro pela soma e dimensionamento de dados por Pareto, que consiste na centralização dos dados na média e a consecutiva divisão de cada variável pela raiz quadrada de seu desvio padrão (Emwas et al., 2018).

A matriz foi submetida à análise exploratória não supervisionada dos dados com o método de análise de componentes principais (PCA) para identificação dos metabólitos dos

grãos de sorgo responsáveis pelas diferenças entre os cultivares. Além disso, a distância euclidiana foi utilizada para realização da análise de grupos hierárquicos (HCA) associada ao mapa de calor a partir das separações observadas pelo PCA.

MetaboAnalyst 6.0 também foi utilizado para a análise de correlação entre os bins e as propriedades das amostras. Para isso, os metadados de tratamento (amostras *in natura* ou desengorduradas) e teor de taninos (alto, intermediário ou baixo) foram classificados como dados categóricos, enquanto o teor de etanol foi tratado como um dado contínuo. A correlação de Spearman foi adotada como medida de correlação nas análises de correlação parcial. O algoritmo de classificação floresta aleatória (*random forest*) foi utilizado para avaliar o desempenho classificatório da metabolômica (Gupta et al., 2023; Pang et al., 2022; Xia & Wishart, 2011).

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Tratamento das amostras: desengorduramento dos grãos

O desengorduramento dos grãos foi realizado objetivando-se avaliar se o teor de óleo presente nos grãos de sorgo influenciaria o rendimento final na obtenção do bioetanol. Assim, um novo grupo de trabalho foi gerado, o das amostras de sorgo desengorduradas (Tabela 2). Para fins de nomenclatura, às amostras desengorduradas foi adicionado o sufixo “\_d” após o nome da cultivar de origem.

Tabela 2. Perda de massa (%) das amostras de sorgo ocasionadas pelo tratamento de desengorduramento dos grãos e teor de extrato hexânico (% m/m)

| Amostra  | Perda de massa pelo desengorduramento (%) | Extrato hexânico (% m/m) |
|----------|---|--------------------------|
| SC084    | 5,12                                      | 2,77                     |
| BRS305   | 1,34                                      | 3,38                     |
| CMSXS180 | 2,03                                      | 2,28                     |
| BR501    | 2,85                                      | 3,31                     |

O desengorduramento dos grãos de sorgo resultou em uma perda média de 2,8% de massa, em percentual de base seca. As amostras BRS305 e BR501 apresentaram maiores teores de lipídeos, contrapondo com a amostra CMSXS180, que apresentou o menor teor (Tabela 2).

### 5.2 Caracterização química das amostras de sorgo

#### 5.2.1 Determinação do teor de compostos fenólicos totais

Duas metodologias utilizando o reagente Folin-Ciocalteu foram empregadas, onde o método, solventes de extração da amostra, a base utilizada e o comprimento de onda foram diferentes entre eles. O Método 1 foi a análise padrão dos laboratórios da Embrapa Milho e Sorgo. Esse método trabalha com extração em meio ácido e grandes volumes de amostra, sendo as leituras realizadas em cubetas e com maior demanda de tempo. O Método 2 segue o protocolo do laboratório CerQBio, que foi desenvolvido para análise de extratos de plantas em placas de 96 poços, onde a massa da amostra e o volume dos reagentes são menores e, considerando-se

triplicatas, 20 amostras podem ser lidas de uma só vez, otimizando a análise. Ambos expressam os resultados em mg de ácido gálico (GA) por grama de amostra seca.

Dentre as amostras analisadas, as cultivares SC084 e BRS305 apresentaram teores mais elevados de compostos fenólicos totais (Figura 7), sendo que a cultivar SC084 é a que possui o maior teor dentre todas as amostras.

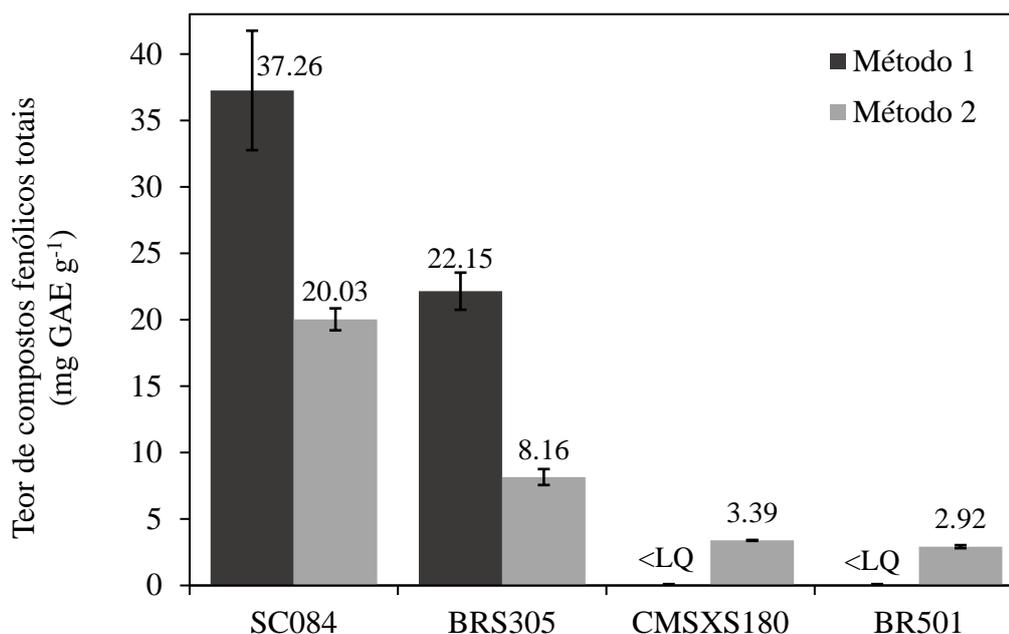


Figura 7. Teor de compostos fenólicos totais determinados por duas metodologias diferentes para as amostras de sorgo *in natura*. GAE (Ácido Gálico Equivalente).

As amostras CMSXS180 e BR501 não apresentaram teores quantificáveis de compostos fenólicos totais de acordo com o método 1, ficando abaixo do limite de quantificação da curva de calibração utilizada ( $LQ = 0,178 \text{ mg GAE g}^{-1} \text{ amostra}$ ), enquanto, pelo método 2, apresentaram pouco menos da metade do teor da cultivar BRS305 e teores quase 6 vezes menores que a cultivar SC084, sendo que para ambas as amostras (CMSXS180 e BR501) não foi observada diferenciação significativa entre os valores obtidos.

A metodologia 1, empregada no laboratório da Embrapa, envolveu a extração dos compostos fenólicos utilizando-se metanol acidificado com leitura da absorção em 600 nm, e mesmo utilizando maior quantidade de amostras se mostrou menos sensível. De acordo com Pérez e colaboradores (2023) compostos fenólicos transferem elétrons aos complexos de ácido fosfomolibdico / fosfotúngstico em uma solução alcalina (pH 10), gerando complexos azuis que são detectados espectroscopicamente a 760nm. A extração em meio ácido pode ter

influenciado no pH final da solução alcalinizada após a adição da solução de etalonamina. A leitura feita em 600 nm também pode ter influenciado na sensibilidade do método. A metodologia 2, realizada no laboratório CerQBio/UFMG, utilizou solução aquosa de metanol (1:1), adição de carbonato a 35%, leitura em 760 nm e maior tempo de reação (20 minutos), requisitos elencados para uma maior acurácia do teste (Pérez et al., 2023).

Embora as amostras de pericarpo pigmentado (vermelho) tenham sido as que apresentaram maiores teores de compostos fenólicos totais, contrapondo com as amostras de pericarpo claro/branco, a existência de pigmentação nos grãos não pode ser considerada um indicativo para a afirmação de que a cultivar contém compostos fenólicos, assim como a pigmentação mais intensa não pode ser considerada indicativo de teores mais elevados. Em estudos anteriores, Rao e colaboradores (2018) relataram amostras de sorgo com pericarpos vermelho e branco que apresentaram baixos teores de compostos fenólicos totais e proantocianidinas, sendo que os valores reportados para as quatro amostras em questão não apresentaram diferença significativa entre si. Em outro trabalho, Awika e colaboradores (2005) investigaram variedades de sorgo onde as amostras com pericarpo preto continham teores mais baixos de compostos fenólicos totais quando comparadas com amostras de pericarpo marrom.

### 5.2.2 Determinação do teor de taninos pelo método Vanilina-HCl

Dentre as amostras analisadas, apenas as cultivares SC084 e BRS305 apresentaram teores de taninos quantificáveis pelo método utilizado (Tabela 3). A cultivar SC084 apresentou o maior teor de taninos, com  $94,89 \pm 2,24$  mg CE g<sup>-1</sup> amostra, seguido da cultivar BRS305, com  $56,46 \pm 0,77$  mg CE g<sup>-1</sup> amostra. As cultivares CMSXS180 e BR501 apresentaram resultado abaixo do limite de quantificação da curva de calibração utilizada (LQ = 1,75 mg CE g<sup>-1</sup> amostra).

Tabela 3. Teor de taninos para as amostras de sorgo *in natura*

| Amostra  | Teor de taninos (mg CE g <sup>-1</sup> amostra) |
|----------|---|
| SC084    | $94,89 \pm 2,24$                                |
| BRS305   | $56,46 \pm 0,77$                                |
| CMSXS180 | <LQ   |
| BR501    | <LQ   |

CE (Equivalente em catequina). LQ (Limite de Quantificação) = 1,75 mg CE g<sup>-1</sup> amostra.

### 5.2.3 Determinação do teor de amido total

Os teores de amido total foram determinados para as quatro cultivares de sorgo em ambas as amostras *in natura*<sup>2</sup> e desengorduradas. Dentre as amostras analisadas (Figura 8), a cultivar BRS305 foi o que apresentou maior teor de amido total tanto para a amostra *in natura* quanto para a amostra desengordurada (67,4% e 69,9%, respectivamente), seguido pela cultivar CMSXS180 (66,6% *in natura* e 67,0% desengordurada). As cultivares com menor teor de amido total foram SC084 (62,9% *in natura* e 64,3% desengordurada) e BR501 (63,4% *in natura* e 63,8% desengordurada). O tratamento de desengorduramento dos grãos de sorgo resultou em um aumento de 4,9% no teor médio de amido das amostras.

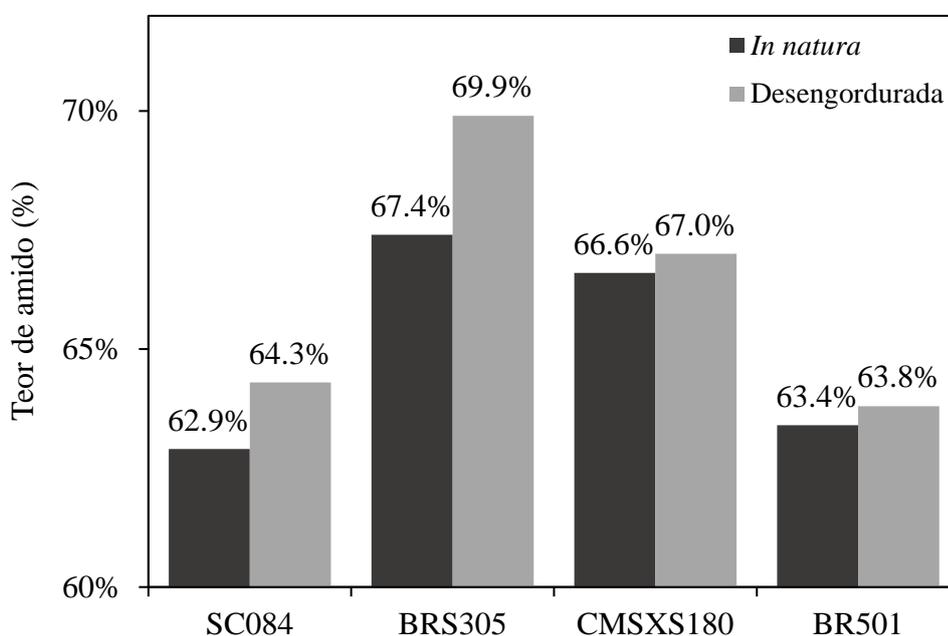


Figura 8. Teor de amido total (%) para as amostras de sorgo *in natura* e desengorduradas.

Uma vez que o amido é convertido em açúcares fermentáveis durante a obtenção do etanol, espera-se que um maior teor de amido contribua para um maior rendimento do processo de fermentação. Sendo assim, pode-se esperar que as amostras desengorduradas apresentem um desempenho melhor na fermentação, dado o aumento médio no teor de amido total quando comparadas com as amostras *in natura*.

Dentre as cultivares *in natura*, analisando apenas o teor de amido total, pode-se esperar que a cultivar BRS305 apresente um maior rendimento no processo de produção de etanol,

<sup>2</sup> Para as amostras *in natura*, o teor de amido não foi determinado em triplicata

seguido pela CMSXS180. As cultivares SC084 e BRS305 apresentaram teores similares de amido, resultando em rendimentos teóricos similares no processo de produção de etanol. No entanto, estudos prévios reportaram a presença de taninos como um possível interferente no processo de produção de etanol, afetando a atividade enzimática durante a conversão do amido e ocasionando a diminuição do rendimento do processo quando comparado a amostras sem taninos em sua composição (Chuck-Hernandez et al., 2012; Foglia et al., 2022). Dessa forma, espera-se que a cultivar BRS305, apesar de apresentar rendimento teórico similar à SC084, se destaque quanto ao rendimento durante o processo fermentativo.

#### 5.2.4 Determinação do perfil de antocianidinas por CLAE

Para avaliar o teor de antocianidinas nas amostras de sorgo, as antocianidinas apigeninidina, 7-metoxi-apigeninidina, luteolinidina, 5-metoxi-luteolinidina e 5,7-dimetoxi-luteolinidina (Figura 5, página 35) foram quantificadas por Cromatografia Líquida de Alta Eficiência (CLAE). Dentre essas, apenas a 5,7-dimetoxi-luteolinidina não foi detectada pelo limite de quantificação do método em nenhuma das cultivares.

A cultivar SC084 foi a que apresentou teores mais elevados das antocianidinas analisadas (Tabela 4). A apigeninidina e seu derivado 7-metoxilado se destacaram com maior concentração nessa amostra, totalizando 83% do teor total de antocianidinas. A cultivar híbrida BRS305 apresentou concentrações mais baixas para as mesmas antocianidinas analisadas. Para essa amostra, a luteolinidina e a sua derivação 5-metoxilada se destacaram, sendo responsáveis por 80% da concentração total nessa amostra. As cultivares CMSXS180 e BR501 não apresentaram teores quantificáveis de antocianidinas em sua composição.

Tabela 4. Teor de antocianidinas, em  $\mu\text{g g}^{-1}$ , determinado por CLAE

| Amostra  | Teor de antocianidinas ( $\mu\text{g g}^{-1}$ ) |                          |                        |                        |                            |
|----------|---|--------------------------|------------------------|------------------------|----------------------------|
|          | Cloreto de luteolinidina                        | Cloreto de apigeninidina | 5-metoxi-luteolinidina | 7-metoxi-apigeninidina | 5,7-dimetoxi-luteolinidina |
| SC084    | 31,85   | 140,37                   | 20,51                  | 124,22                 | ND                         |
| BRS305   | 8,83  | 0,76                     | 13,57                  | 4,70                   | ND                         |
| CMSXS180 | ND  | ND                       | ND                     | ND                     | ND                         |
| BR501    | ND  | ND                       | ND                     | ND                     | ND                         |

ND - Não detectado no limite de quantificação do método.

### 5.2.6 Determinação da composição centesimal

A Análise Centesimal estuda os alimentos, mais especificamente os valores nutricionais de um produto. Através desta análise é possível identificar, por exemplo, o percentual de umidade, fibras, lipídios e proteínas presentes (Tabela 5).

O teor de umidade determina o percentual de massa de água contida na amostra com relação à sua massa de matéria seca. A análise foi realizada para as amostras *in natura* e desengorduradas de forma a possibilitar a conversão de todos os resultados para base seca, possibilitando que a massa de água presente na amostra não seja considerada um fator interferente nos resultados finais obtidos tanto para a caracterização e composição das amostras de sorgo como para o processo de fermentação e produção do etanol.

O extrato etéreo é a soma das substâncias extraídas pelo éter que, pela metodologia aplicada nos laboratórios da Embrapa (AOCS, 2005; Gomes & Simeone, 2012), são extraídas sob alta pressão (não informada) e aquecimento a 90 °C em um extrator de gordura Ankom modelo XT10. Para o procedimento realizado durante o processo de desengorduramento dos grãos, foi obtido um extrato hexânico, extraído sob temperaturas mais brandas (40°C) para não correr o risco de alterar a composição dos grãos de sorgo pela degradação térmica dos metabólitos. O processo de desengorduramento resultou em uma diminuição de 71 a 89% no teor de lipídeos presente nas amostras, o que pode ser observado pelos teores de extrato etéreo das amostras *in natura* comparados às respectivas amostras desengorduradas.

Tabela 5. Composição centesimal das amostras de sorgo *in natura* e desengorduradas (% m/m) para as diferentes cultivares

| Amostra    | Extrato Etéreo | Matéria mineral | Proteína Bruta | Umidade    |
|------------|----------------|-----------------|----------------|------------|
| SC084      | 3,3 ± 0,1      | 1,9 ± 0,0       | 14,6 ± 0,2     | 11,4 ± 0,1 |
| BRS305     | 3,6 ± 0,1      | 1,6 ± 0,0       | 14,4 ± 0,2     | 10,7 ± 0,1 |
| CMSXS180   | 2,4 ± 0,1      | 1,6 ± 0,0       | 13,8 ± 0,2     | 10,9 ± 0,2 |
| BR501      | 3,8 ± 0,1      | 1,6 ± 0,0       | 13,0 ± 0,1     | 10,5 ± 0,1 |
| SC084_d    | 0,9 ± 0        | 1,7 ± 0,0       | 13,1 ± 0,2     | 6,0 ± 0,2  |
| BRS305_d   | 0,4 ± 0        | 1,5 ± 0,1       | 11,1 ± 0,1     | 6,5 ± 0,1  |
| CMSXS180_d | 0,6 ± 0        | 1,3 ± 0,0       | 13,0 ± 0,1     | 6,8 ± 0,1  |

---

|         |         |           |            |           |
|---------|---------|-----------|------------|-----------|
| BR501_d | 1,1 ± 0 | 1,8 ± 0,0 | 11,9 ± 0,1 | 7,4 ± 0,1 |
|---------|---------|-----------|------------|-----------|

---

A matéria mineral influencia na proporção de amido no grão de forma que quanto maior seu teor, menos o teor de amido (Chao et al., 2024). Para que o amido dos grãos de sorgo esteja disponível, é necessário que ocorra a digestão das frações proteicas, de forma que o teor de proteína bruta pode ser um interferente na fermentação (Rooney & Miller, 1981). Os lipídeos podem formar complexos com amido dificultando a sacarificação. A cultivar SC084 apresenta o maior teor de proteína e alto teor de lipídeos. Dessa forma, espera-se que essa cultivar apresente menor rendimento no processo de produção de etanol, dados o menor teor de amido e maiores teores de proteínas, lipídeos e compostos fenólicos/taninos.

### 5.3 Caracterizações espectroscópicas

#### 5.3.1 Determinação do perfil no ultravioleta

O perfil no espectro na região do ultravioleta-visível (Figura 9) é condizente com a presença de compostos fenólicos, provavelmente da classe dos flavonoides, e também com a presença em menor quantidade das antocianidinas apigeninidina e luteolinidina, que são 3-desoxiantocianidinas. As 3-desoxiantocianidinas, diferentemente das outras antocianidinas, apresentam o máximo de absorção em regiões do ultravioleta próximo (320, 400 e 489 nm) e não do visível (516 nm) (Awika et al., 2004). A ausência da hidroxila na posição 3 faz das 3-deoxiantocianidinas mais estáveis em solução que as outras cianidinas.

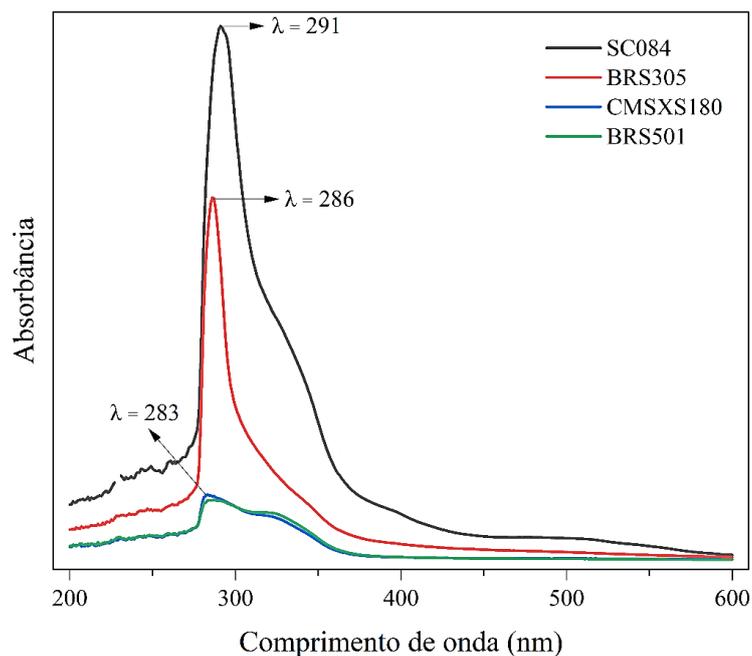


Figura 9. Perfil do espectro no ultravioleta-visível (200 a 600 nm) para as amostras de sorgo, indicando os comprimentos de onda máximos de cada amostra.

### 5.3.2 Determinação da atividade antioxidante

Para as amostras de sorgo analisadas (Tabela 6), a cultivar SC084 foi o que apresentou maior atividade antioxidante, seguido pelo híbrido BRS305. As amostras CMSXS180 e BR501 apresentaram baixa atividade antioxidante.

As duas amostras com alta atividade antioxidante também apresentaram teores elevados de compostos fenólicos. Os compostos fenólicos são substâncias antioxidantes capazes de interagir com espécies radicalares, sendo consumidos durante essa reação (Angelo & Jorge, 2007). Dessa forma, o resultado obtido está de acordo com o esperado, uma vez que o teor de compostos fenólicos está coerente com a atividade antioxidante apresentada pelas amostras.

Tabela 6. Atividade antioxidante das amostras de sorgo *in natura*, expressa em equivalentes de Trolox ( $\mu\text{mol TE g}^{-1}$  amostra)

| Amostra  | Atividade antioxidante<br>( $\mu\text{mol TE.g}^{-1}$ amostra) |
|----------|--|
| SC084    | 524,46 $\pm$ 41,48   |
| BRS305   | 298,17 $\pm$ 25,20   |
| CMSXS180 | 36,31 $\pm$ 0,21   |

---

|       |              |
|-------|--------------|
| BR501 | 23,72 ± 0,96 |
|-------|--------------|

---

TE (Trolox Equivalente).

#### 5.4 Determinação do perfil metabólico por RMN

Ressonância magnética nuclear tem sido extensivamente usada nos diferentes campos da ciência, proporcionando uma visão abrangente do perfil metabólico de diferentes matrizes. Mesmo sendo uma técnica menos sensível que a espectrometria de massas ela oferece a vantagem de não ser uma técnica destrutiva, não gera subprodutos que possam se acumular no aparelho e demandar limpeza constante, é altamente reprodutível e permite a identificação simultânea de vários compostos presentes em um extrato. A análise de uma matriz complexa requer a análise cuidadosa do espectro em uma dimensão, bem como de mapas de contorno em duas dimensões.

A RMN tem sido também uma das técnicas analíticas mais utilizada na análise metabolômica. Caracteriza-se por uma abordagem holística que oferece uma maneira mais abrangente de analisar e correlacionar o metaboloma com condições de estresse biótico ou abiótico e tem sido utilizada em diversas áreas da ciência com diferentes propósitos. Entre eles, a metabolômica tem sido usada para examinar como a diversidade genética e as condições ambientais afetam os fenótipos em plantas. Esta técnica já foi utilizada em análise de sorgo proporcionando informações acerca dos caminhos metabólicos para a biossíntese de flavonoides em sementes de *Sorghum bicolor* branco, vermelho e preto (Zhou et al., 2022). Mandrone e colaboradores (2021) acoplaram a pesquisa agrometereológica à metabolômica não objetivada para estudar a variação do metaboloma de *Sorghum bicolor* (L.) Moench durante o desenvolvimento da cultura, em resposta a fatores ambientais e antrópicos.

Neste trabalho a metabolômica por RMN de  $^1\text{H}$  foi utilizada como técnica para diferenciar o perfil metabólico de quatro genótipos de sorgo produzidos pela Embrapa para a produção de etanol.

A diferença nos espectros de RMN de  $^1\text{H}$  dos diferentes cultivares analisados é visível, principalmente para a amostra SC084 (Figura 10). Os espectros mostraram sobreposição de sinais com intensidades variáveis entre  $\delta 0,85$  e  $8,50$ , indicando a presença de vários compostos e múltiplos grupos funcionais. A maioria dos sinais foi observada na região característica de hidrogênios carbinólicos ( $\delta 3,0-4,2$ ) para todas as cultivares, seguida pela região de hidrogênios

de alcanos ( $\delta$  0,0-1,6). A cultivar SC084 apresentou as maiores intensidades de sinais na região de hidrogênios aromáticos ( $\delta$  6,4-8,5). Este resultado corrobora os resultados discutidos anteriormente, onde a presença de compostos fenólicos foi quantificada majoritariamente neste cultivar.

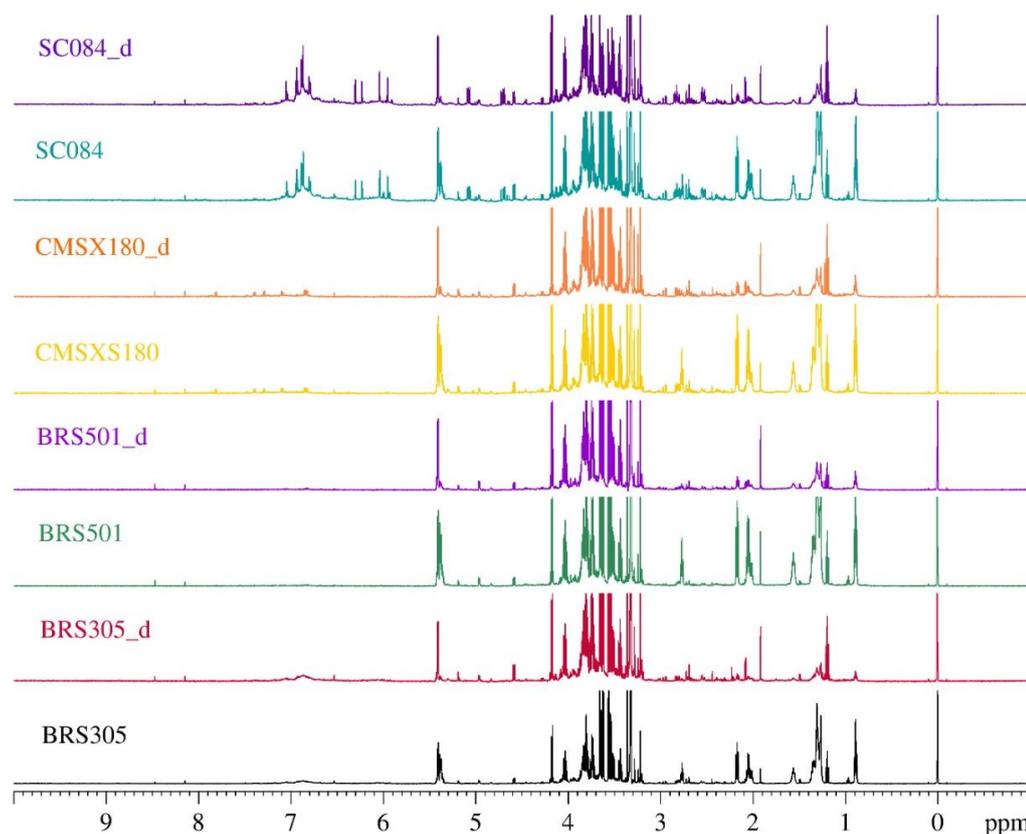


Figura 10. a) Espectros de RMN de  $^1\text{H}$  dos extratos dos grãos de sorgo *in natura* e desengordurados (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0).

Em amostras de matrizes orgânicas encontra-se presente um número significativo de compostos, obtendo-se uma grande quantidade de dados. Para uma interpretação a este volume de dados, ferramentas precisam ser empregadas. Análises quimiométricas não supervisionadas foram realizadas utilizando-se o *software* MetaboAnalyst a partir dos dados espectrais de RMN de  $^1\text{H}$ , sendo estas: Análise de Componentes Principais (PCA), Agrupamento Hierárquico de Clusters (HCA) e a construção de mapas de calor (*heatmaps*).

Pela PCA, as componentes principais 1 e 2 explicam mais de 84% da variância total, sendo 51,7% em PC1 e 32,5% em PC2. O gráfico de escores destas duas componentes mostrou que foi possível diferenciar grãos desengordurados e *in natura* (Figura 11), sendo possível perceber a tendência de quatro agrupamentos. Pode-se perceber um agrupamento referente aos

grãos desengordurados, todos dispostos em PC1 positiva, contrapondo-se aos grãos *in natura* em PC1 negativa. Além disso, percebe-se que a cultivar SC084 se destacou com agrupamentos em PC2 positiva para ambas as amostras *in natura* e desengordurada, enquanto as cultivares BRS305, CMSXS180 e BR501 apresentaram agrupamento em PC2 negativa.

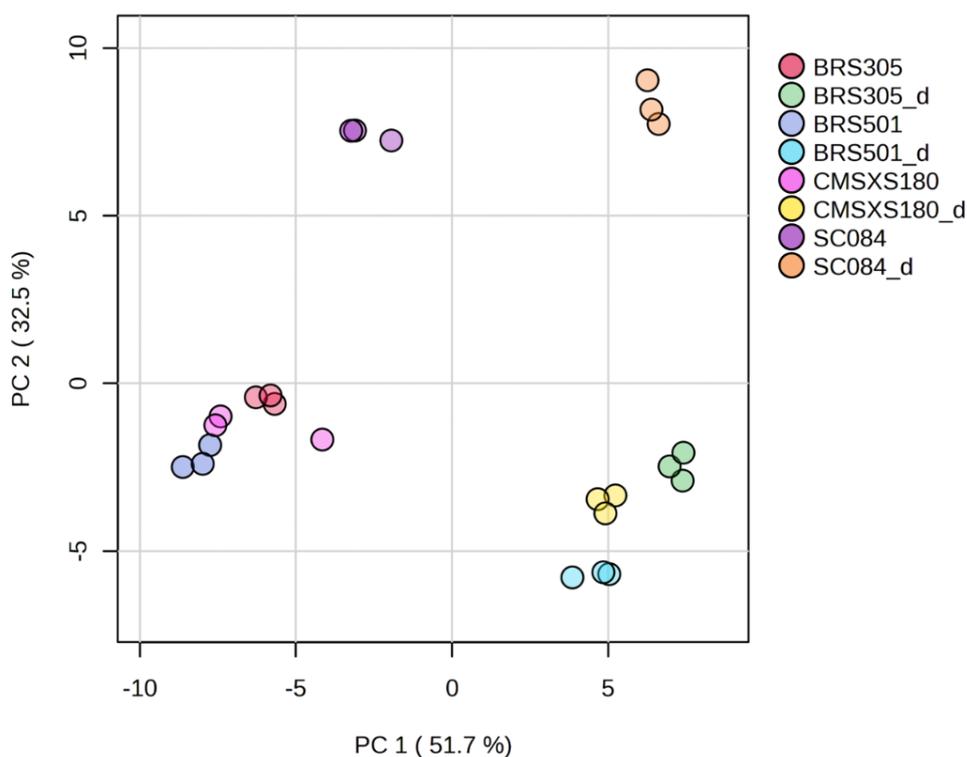


Figura 11. Gráfico de *scores* das componentes principais PC1 (51,7%) e PC2 (32,5%) calculadas a partir dos dados espectrais de RMN de  $^1\text{H}$  das amostras de sorgo *in natura* e desengordurados, construído com o auxílio do *software* MetaboAnalyst 6.0.

Esses agrupamentos mostram que as amostras da cultivar SC084 apresentaram sinais e/ou regiões no espectro de RMN de  $^1\text{H}$  correspondentes a compostos que as demais amostras não possuem, ou possuem em quantidades bem inferiores, e que estes não são extraídos pelo hexano no processo de desengorduramento dos grãos, uma vez que as tendências de separação dos agrupamentos se mantêm as mesmas para as amostras *in natura* e desengorduradas.

Os principais deslocamentos químicos que contribuem para a separação apresentada no gráfico de *scores* (Figura 11) podem ser visualizados no gráfico de peso das componentes PC1 e PC2 (Figura 12).

Na análise do gráfico de *loadings*, é possível identificar quais deslocamentos químicos apresentam maior ou menor contribuição para explicar cada componente principal. Variáveis

mais positivas possuem maior influência no lado positivo da componente principal e variáveis negativas, no lado negativo. As variáveis próximas ao eixo zero, em contrapartida, não apresentam grande influência no modelo da componente principal.

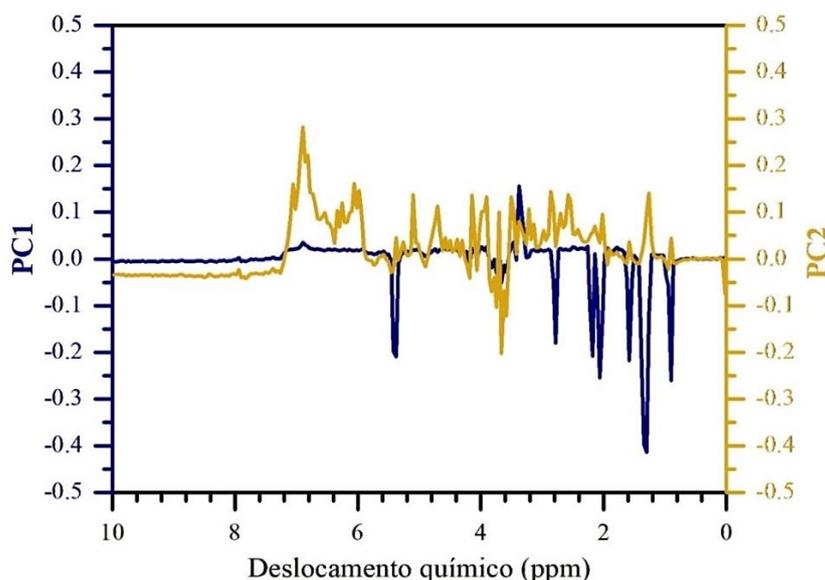


Figura 12. Gráfico de *loadings* das componentes PC1 e PC2, obtido a partir dos dados espectrais de RMN de  $^1\text{H}$  das amostras de sorgo *in natura* e desengordurados.

Pelo gráfico de pesos em PC2 positivo pode-se perceber que os bins centrados em  $\delta$  7,06-5,98; 5,18; 4,70; 4,14; 3,90; 3,70; 3,50-3,14; 2,86 e 2,54; 1,26 e 0,90 apresentaram maior intensidade onde os scores foram positivos, ou seja, para a cultivar SC084. Por outro lado, os bins em  $\delta$  5,50; 5,34; 4,22; 4,06; 3,82; 3,74; 3,66; 3,62; 1,94; 1,62 e 0,94 foram mais intensos na explicação do modelo que inclui as amostras agrupadas em PC2 negativa, sendo essas as cultivares BRS305, CMSXS180 e BR501. PC1 fez a separação das amostras *in natura* daquelas submetidas ao tratamento de desengorduramento. Pelo gráfico de *loadings* foi possível observar que os bins  $\delta$  5,38; 2,78; 2,18; 2,06; 1,58; 1,3 e 0,9 se destacam como maior contribuição influenciando a explicação dos agrupamentos apresentados em PC1 negativa.

A cultivar SC084, que apresentou agrupamento distante das demais, é conhecida como uma linhagem com elevado teor de tanino/compostos fenólicos e mostrou um perfil metabólico no espectro de RMN de  $^1\text{H}$  condizente com a informação e totalmente diferenciado do perfil apresentado pelas demais amostras. Partindo-se dessa informação e considerando os resultados do teor de taninos (Tabela 3, página 46) e fenólicos totais (Figura 7, página 45), procurou-se nos espectros de RMN de  $^1\text{H}$  sinais compatíveis com taninos condensados e flavonoides.

Tendo a catequina como ponto de partida, esperava-se sinais na região de aromáticos, onde comumente dois sinais com constante de acoplamento ( $J$ ) meta são esperados para o anel A na faixa de  $\delta$  5,80-6,10 referente a H-6, e  $\delta$  5,60-5,80 referente a H-8. Para o anel A de flavonoides estes sinais se encontram um pouco mais desprotegidos sendo observados em  $\delta$  6,20-6,30 e  $\delta$  6,30-6,90. Estes sinais puderam ser claramente observados nos espectros de RMN de  $^1\text{H}$  de SC084 (Figura 13) em  $\delta$  6,23 (d,  $J = 2,24$  Hz) e  $\delta$  6,30 (d,  $J = 2,20$  Hz).

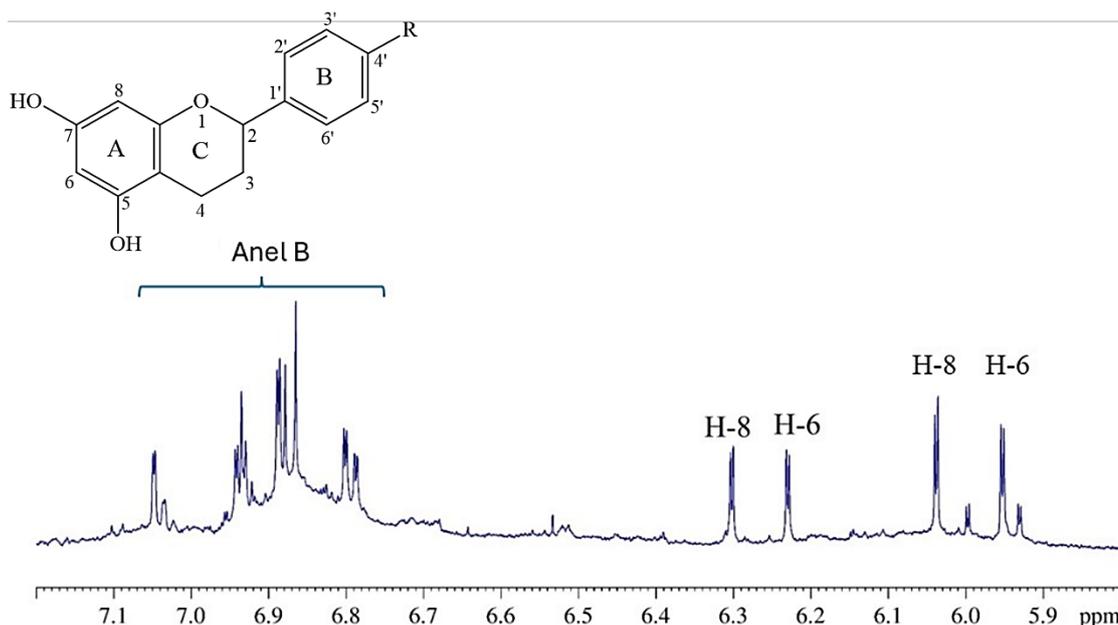


Figura 13. Espectro de RMN de  $^1\text{H}$  da amostra SC084 *in natura* (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0). Expansão em  $\delta$  5,90-7,10 ppm.

Pelo espectro  $J$ -resolvido e pelo mapa de contornos COSY pode-se confirmar o acoplamento de meta entre eles (Figura 14). Para o anel B das catequinas/flavonoides, três sinais são esperados na faixa de  $\delta$  6,60-6,90. No espectro podemos observar um acoplamento tipo AA'B em  $\delta$  6,80 ( $J = 8,50; 2,08$  Hz), 6,94 e 6,88.

Três tipos de flavonoides são encontrados em grande quantidade em diferentes genótipos de sorgo: 3-desoxiantocianidinas, flavonas e flavanonas (Taleon et al., 2012). Estudo metabolômico realizado com grãos de *Sorghum bicolor* utilizando CLAE-EM-EM-EM identificaram 171 flavonoides em três genótipos diferentes (Zhou et al., 2022).

Proantocianidinas, (+)-catequina, procianidina B, taxifolina, taxifolina-7-O- $\beta$ -glicosídeo, narigenina, eriodictiol e seu derivado 5-O- $\beta$ -glicosídeo já foram isolados em sorgo húngariano (Gujer et al., 1986). Os resultados deste estudo estão alinhados com pesquisas

anteriores e destacam a necessidade de novas investigações para identificar os flavonoides mais abundantes nas amostras de sorgo analisadas.

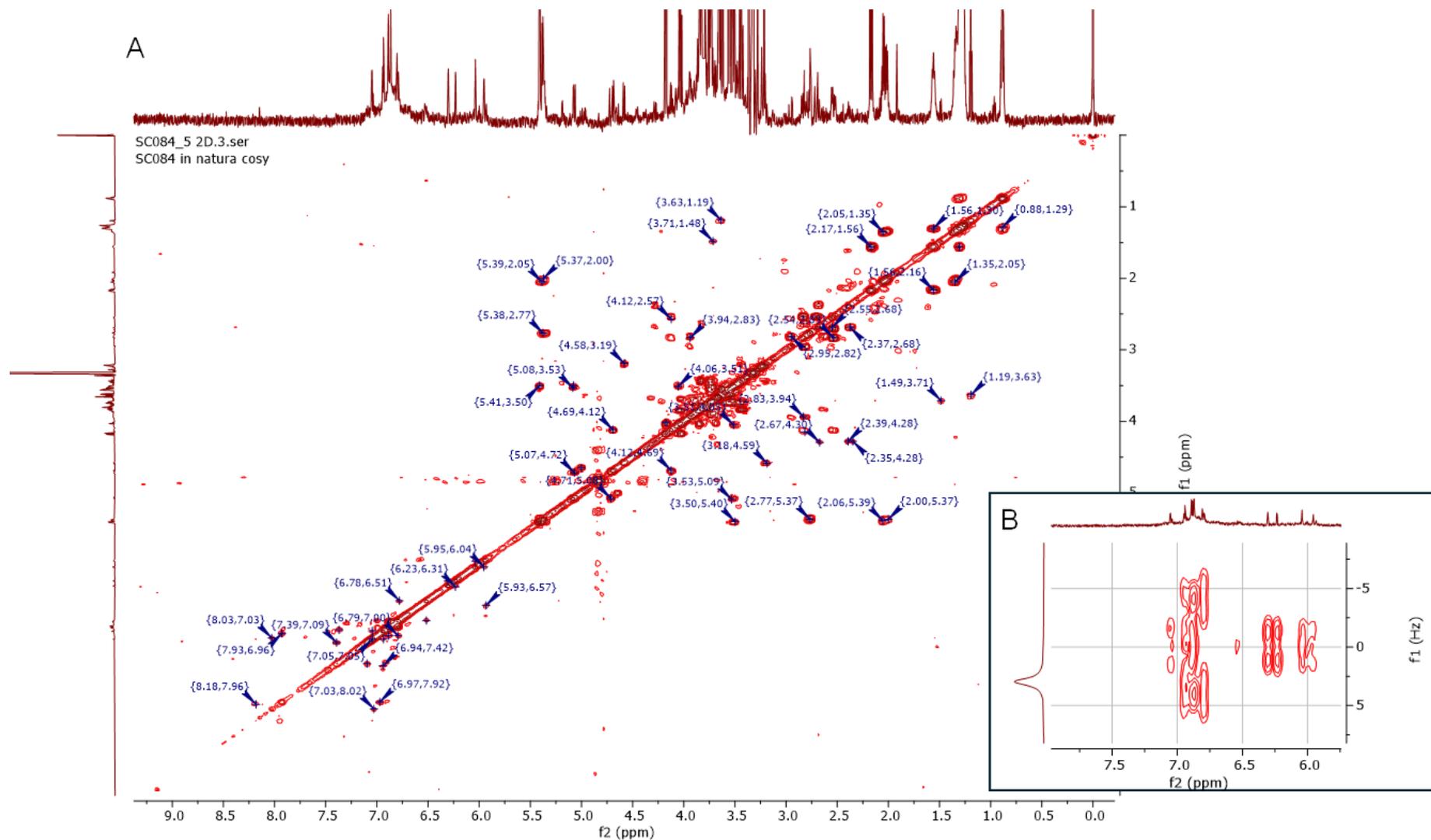


Figura 14. Mapa de contornos COSY (A) e Espectro *J*-resolvido (B) do grão de sorgo da cultivar SC084 (600 MHz, metanol-*d*<sub>4</sub> e tampão KH<sub>2</sub>PO<sub>4</sub> em D<sub>2</sub>O com 0,01% TSP-*d*<sub>4</sub> (m/m), pH 6,0) com expansão em  $\delta$  5,80-7,20.

Além destes sinais relacionados aos derivados flavonoidicos, outros compostos e/ou classe de compostos puderam ser identificados no espectro por comparação do deslocamento químico observado com aqueles descritos na literatura (Tabela 7). Entre os sinais foram encontrados sinais de hidrogênios característicos de compostos fenólicos, de hidrogênios anoméricos de açúcares, de H- $\alpha$  de aminoácidos, ácidos orgânicos e ácidos graxos.

Tabela 7. Atribuições de metabólitos identificados no sorgo pelo perfil no RMN, com deslocamentos químicos ( $\delta$ , ppm), padrão de divisão e constantes de acoplamento ( $J$ , Hz)

| Grupo            | Metabólito                      | Sinal característico no RMN de $^1\text{H}$   |
|------------------|---------------------------------|---|
| Ácidos graxos    | Ácido linoleico*                | $\delta$ 2,75 (m) <sup>1</sup>  |
|                  | Ácido linolênico**              | $\delta$ 0,98 (t, $J = 6,91$ Hz) <sup>1,5</sup>   |
|                  | Ácido oleico                    | $\delta$ 2,02 (m) <sup>1,5</sup>  |
| Ácidos orgânicos | Ácido fórmico                   | $\delta$ 8,47 (s) <sup>2,4</sup>  |
|                  | Ácido fumárico                  | $\delta$ 6,54 (s) <sup>2</sup>  |
| Açúcar           | $\alpha$ -Glicose               | $\delta$ 5,19 (d, $J = 3,91$ Hz) <sup>2,3</sup>   |
|                  | $\beta$ -Glicose                | $\delta$ 4,58 (d, $J = 7,84$ Hz) <sup>2,3</sup>   |
|                  | Sacarose                        | $\delta$ 5,41 (d, $J = 3,67$ Hz); $\delta$ 4,17 (d, $J = 8,77$ Hz) <sup>2,3</sup>             |
|                  | Alanina                         | $\delta$ 1,49 (d, $J = 7,39$ Hz) <sup>2,3</sup>   |
| Aminoácido       | Glutamato                       | $\delta$ 2,09 (m) <sup>2</sup>  |
|                  | Glutamina                       | $\delta$ 2,16 (m) <sup>2</sup>  |
|                  | Tirosina                        | $\delta$ 7,09 (H-2, H-6, $J = 8,89$ Hz); $\delta$ 6,80 (H-3, H-5, $J = 8,68$ Hz) <sup>4</sup> |
| Ácidos orgânicos | Treonina                        | $\delta$ 1,32 (d, $J = 6,94$ Hz) <sup>2</sup>   |
|                  | Ácido fórmico                   | $\delta$ 8,47 (s) <sup>2,4</sup>  |
|                  | Ácido fumárico                  | $\delta$ 6,54 (s) <sup>2</sup>  |
| Flavonoides      | Ácido gálico                    | $\delta$ 7,04 (s) <sup>4</sup>  |
|                  | Ácido <i>p</i> -hidroxibenzoico | $\delta$ 7,92 (d, $J = 7,92$ Hz); $\delta$ 6,80 (d, $J = 7,88$ Hz) <sup>4</sup>               |
| Outros           | Colina                          | $\delta$ 3,22 (s) <sup>2</sup>  |
|                  | Acetato                         | $\delta$ 1,92 (s) <sup>2,3</sup>  |

\* deslocamento característico de ácidos graxos poli-insaturados em n-6

\*\* deslocamento característico de ácidos graxos poli-insaturados em n-3

1. Barison et al., 2010; 2. Kim et al., 2010; 3. Mandrone et al., 2021; 4. Pimenta et al., 2013; 5. Shiao & Shiao, 1989

Os dados provenientes do espectro de RMN de  $^1\text{H}$  também foram utilizados para a construção de um dendrograma de Análise Hierárquica de Clusters (Figura 15). O HCA agrupa as amostras de acordo com as suas similaridades, tendo como resultado uma árvore binária que representa possíveis divisões dos dados em grupos (*clusters*). O processo de construção do HCA é baseado em uma matriz de distâncias onde os elementos com maior proximidade entre as divisões na árvore são considerados mais correlatos entre si (Maraschin et al., 2017). No dendrograma das amostras de sorgo pôde-se constatar uma boa separação do grupo correspondente a cultivar SC084, para ambas as amostras *in natura* e desengordurada, apresentando distanciamento dos demais grupos. Essa segregação corrobora com a PCA reforçando a presença de metabólitos diferenciados nesse cultivar. Na distância euclidiana próximo a 12 podemos observar que também é possível diferenciar BRS305 de BR501 e CMSX180.

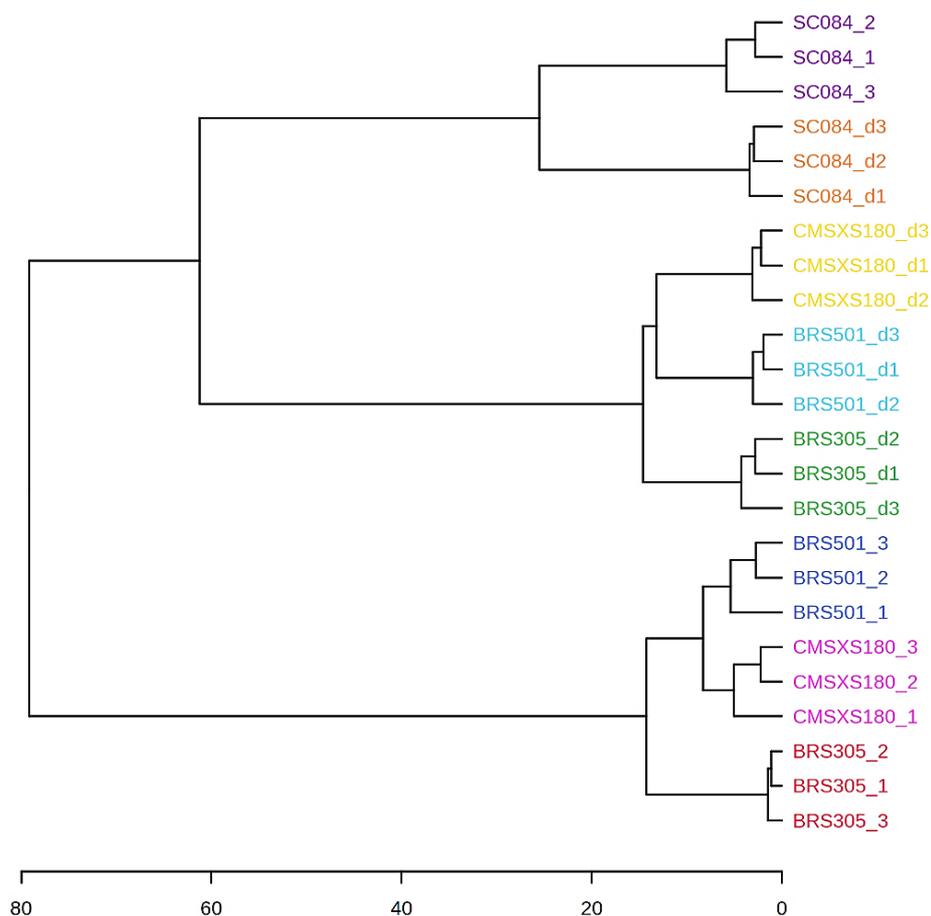


Figura 15. Análise Hierárquica de Clusters (HCA) dos dados de RMN de  $^1\text{H}$  das amostras de sorgo construído com o auxílio do *software* MetaboAnalyst.

Uma forma de visualizar o agrupamento hierárquico e identificar os elementos de agregação que evidenciam as diferenças entre os cultivares de forma simultânea são os *heatmaps* - mapas de calor (Figura 16). O mapa de calor mostra as correlações entre os sinais de deslocamentos químicos dos dados espectrais de RMN de  $^1\text{H}$ , aqui apresentados na forma de bins, e as amostras, evidenciando o grau de importância de cada correlação (Maraschin et al., 2017).

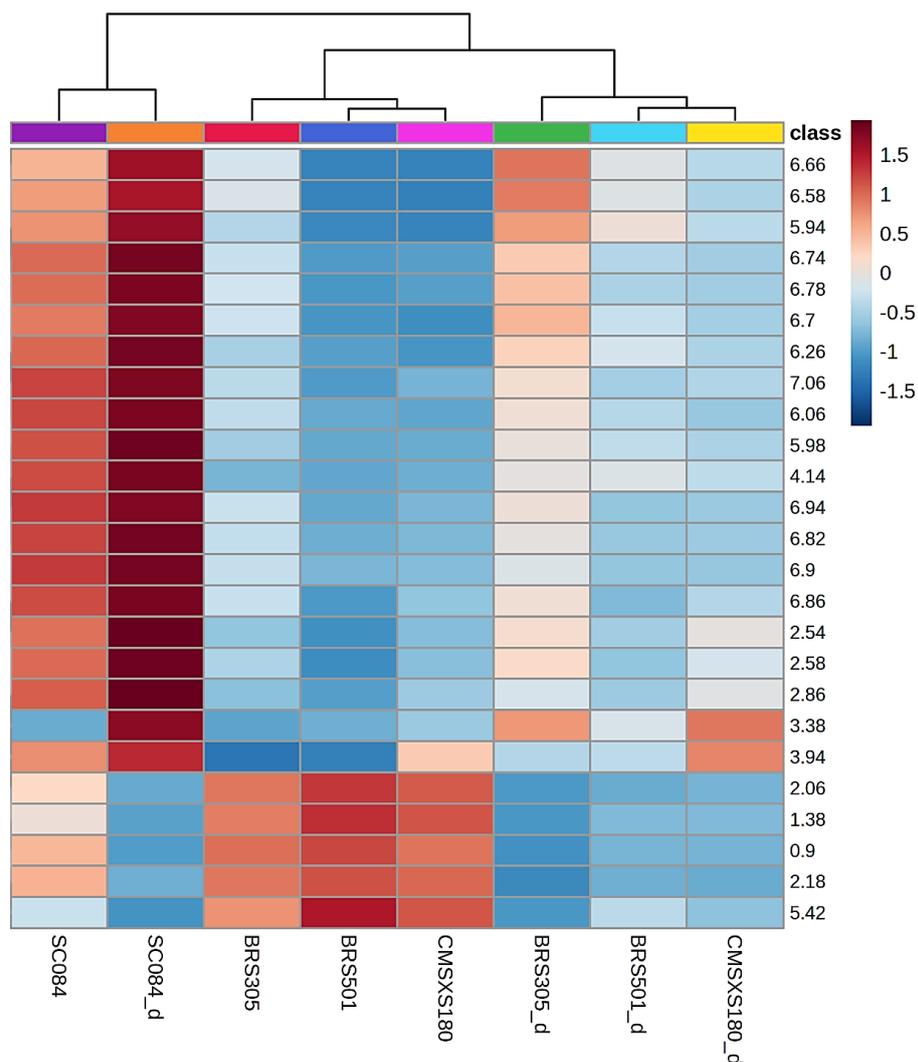


Figura 16. Mapa de calor dos dados espectrais de RMN de  $^1\text{H}$  das amostras de sorgo *in natura* e desengorduradas em estudo construído com o auxílio do software *MetaboAnalyst*, onde as linhas representam os sinais de ressonância (ppm) e as colunas representam as amostras agrupadas de acordo com a Análise Hierárquica de Clusters (HCA) dos dados. Valores em vermelho com cores mais escuras indicam maior importância dos valores e cores em azul indicam valores menos importantes.

A partir do mapa de calor construído, foi possível estabelecer os sinais de maior importância para a diferenciação das amostras analisadas. Pode-se perceber que os sinais de hidrogênio na região de aromáticos são de alto grau de relevância para as amostras da cultivar SC084, evidenciando que o teor de compostos fenólicos nessa cultivar é um fator de diferenciação entre ele e as demais. Ainda que a cultivar BRS305 também possua compostos fenólicos em sua composição, o teor destes não se equivale àquele da linhagem SC084, e somente se evidencia nas amostras desengorduradas. Outro metabólito diferenciador entre os genótipos foram os ácidos graxos insaturados que parecem estar mais abundantes em BR501 e CMSXS180  $\delta$  0,88 ( $\delta_c$  16,36); 1,30 ( $\delta_c$  31,74); 2,05 ( $\delta_c$  29,54); 2,16 ( $\delta_c$  40,29) e 5,40 ( $\delta_c$  95,22), sendo também observados, em menor proporção, na amostra BRS305.

### 5.5 Obtenção do etanol

Uma vez constatado que os perfis metabólicos entre as cultivares apresentavam diferenças, partiu-se para o estudo sobre a possibilidade de correlação deste perfil metabólico à eficiência na produção de etanol. A obtenção do etanol a partir do sorgo granífero foi realizada por via seca baseando-se na metodologia utilizada na Embrapa para produção de etanol a partir de grãos de milho (Cardoso et al., 2020). Durante o processo, foi realizada a hidrólise enzimática em duas etapas, sendo a primeira de liquefação e a segunda de sacarificação. Durante a liquefação, a enzima  $\alpha$ -amilase atuou na quebra das ligações  $\alpha$ -1,4 do amido presente nas amostras, originando maltodextrinas e maltoses como produtos parcialmente hidrolisados (Figura 17). Em seguida, na etapa de sacarificação, a enzima amiloglicosidase atua sobre as ligações  $\alpha$ -1,4 das extremidades não redutoras e sobre as ligações glicosídicas  $\alpha$ -1,6 da amilopectina, resultando na obtenção de glicose (van der Maarel et al., 2002). A glicose obtida pela sacarificação foi, então, fermentada pela levedura *Saccharomyces cerevisiae*.

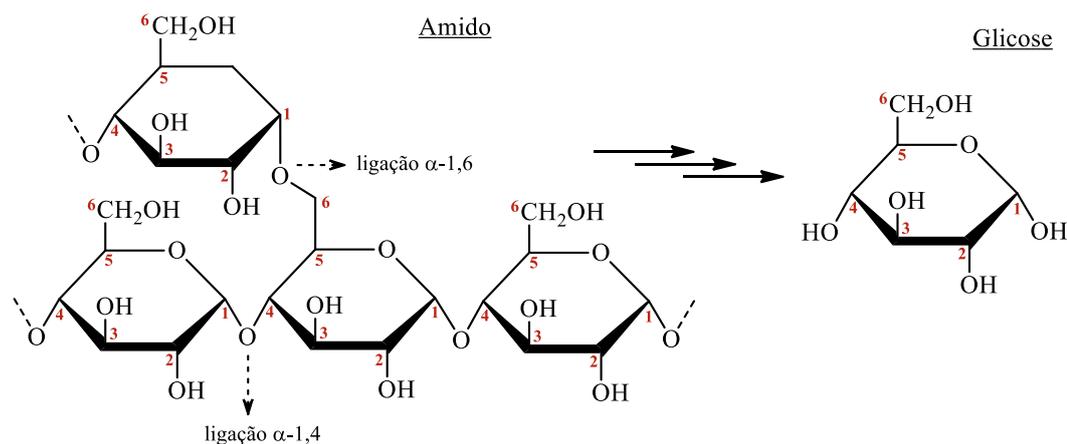


Figura 17. Estruturas das moléculas de amido e glicose, indicando as ligações  $\alpha$ -1,4 e  $\alpha$ -1,6 presentes na molécula de amido.

Na obtenção do etanol foi utilizado o processo de sacarificação e fermentação simultâneas (SSF – *Simultaneous Saccharification and Fermentation*), com a adição da amiloglucosidase e da levedura ao mesmo tempo, em uma única etapa. A SSF é possível com a utilização de enzimas que atuem na sacarificação em temperaturas semelhantes às leveduras da fermentação (32-35°C) e sua aplicação é interessante porque reduz o tempo total do processo de obtenção do etanol, além de reduzir a incidência de inibição da enzima pela glicose produzida, uma vez que esta é formada e fermentada simultaneamente. Além disso, a menor concentração inicial de glicose na solução resulta em um menor risco de contaminação do processo (Pinaki et al., 2015; Sánchez & Cardona, 2008).

Cardoso e colaboradores (2020) reportaram, para as amostras de grãos de milho, rendimento esperado de 0,253 a 0,500 mL de etanol anidro para cada grama de milho, com coeficiente de variação de até 10% entre as triplicatas. A fermentação das amostras de sorgo resultou em teores de etanol entre 0,351 e 0,423 mL por grama de sorgo em base seca (Figura 18), com coeficiente de variação máximo de 4% entre as triplicatas.

O desengorduramento dos grãos resultou no aumento médio no teor de etanol produzido, sendo que as amostras sem a presença de compostos fenólicos em sua composição apresentaram um aumento mais expressivo (CMSXS180: 5,7%; e BR501: 4,9%) que as amostras contendo compostos fenólicos (SC084: 2,6%; BRS305: 3,2%).

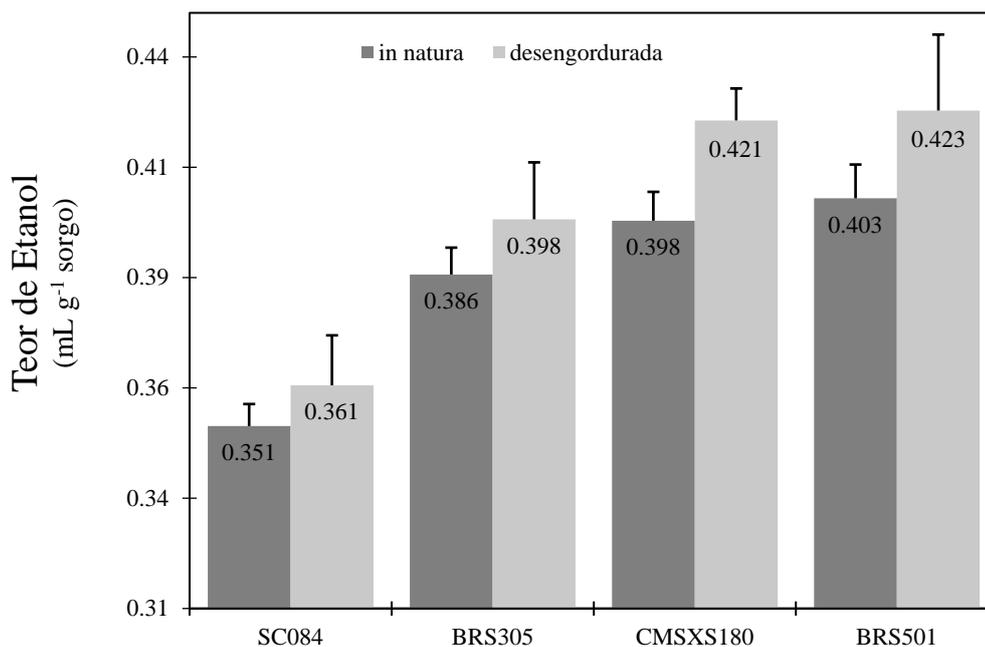


Figura 18. Teores de etanol, em mL g<sup>-1</sup> de sorgo, obtidos a partir da fermentação das amostras *in natura* e desengorduradas.

As cultivares CMSXS180 e BR501, ambas com baixo teor de compostos fenólicos em sua composição, foram as que resultaram em teores mais elevados de etanol (Tabela 9, apêndice, página 82). SC084 foi a cultivar que resultou no menor teor de etanol dentre todos, o que pode ser atribuído ao seu alto teor de compostos fenólicos, lipídeos e proteínas, que foram previamente reportados como possíveis interferentes no processo enzimático de conversão do amido a glicose (Chao et al., 2024; Chuck-Hernandez et al., 2012; Foglia et al., 2022).

Como as amostras de cultivares e com tratamentos diferentes possuem teores de amido diferentes (Figura 8, página 26), os valores teóricos de rendimento devem ser calculados com a finalidade de comparar a eficiência da fermentação. O cálculo é feito considerando-se a estequiometria da reação, onde 1 mol de amido produz 2 mol de etanol (Tabela 8).

A cultivar BRS305 foi a amostra com maior teor de amido total para ambas as amostras *in natura* e desengordurada e, conseqüentemente, produziria um maior rendimento no processo de produção de etanol (Tabela 8). No entanto, a amostra BR501, que apresentou o segundo menor teor de amido, foi a que apresentou o maior rendimento no processo de produção do etanol. A influência da presença de compostos fenólicos e lipídeos pode ser destacada na cultivar BRS305 que, apesar de apresentar teor em taninos muito inferior à SC084, teve a conversão de amido a glicose prejudicada, diminuindo o rendimento da reação. Dessa forma,

pode-se relacionar a presença de compostos fenólicos e lipídeos à interferência no processo de produção do etanol mesmo quando em baixas quantidades na amostra. Assim, a cultivar SC084, como esperado, apresentou o menor rendimento para a reação, tanto para a amostra *in natura* quanto desengordurada.

Tabela 8. Rendimento do processo de fermentação para as amostras de sorgo *in natura* e desengorduradas

| Amostra    | Produção teórica de etanol (mL) | Rendimento do processo (%) |
|------------|---------------------------------|----------------------------|
| SC084      | 0.404                           | 86.93                      |
| SC084_d    | 0.430                           | 83.82                      |
| BRS305     | 0.432                           | 89.29                      |
| BRS305_d   | 0.464                           | 85.84                      |
| CMSXS180   | 0.431                           | 92.31                      |
| CMSXS180_d | 0.443                           | 94.93                      |
| BR501      | 0.406                           | 99.29                      |
| BR501_d    | 0.428                           | 98.72                      |

A não obtenção do teor de amido para todas as 24 amostras inviabilizou a aplicação da Análise por Mínimo Quadrados Parciais (PLS), onde o rendimento do processo fermentativo pudesse ser diretamente correlacionado ao perfil metabólico das diferentes cultivares. Além do teor de etanol, também foram quantificados os teores de glicose e glicerol (Tabela 10, apêndice A, página 82). Essas quantificações foram feitas para determinar a evolução do processo e garantir que a glicose, obtida pelo processo enzimático, foi devidamente convertida pelas leveduras durante a fermentação. Os valores possuem altos desvios padrões para as triplicatas devido ao fato de ser uma quantificação residual e proveniente de um processo biológico, que é mais dificilmente controlado do que uma reação química em escala laboratorial sem o envolvimento de leveduras.

Na tentativa de ilustrar a correlação do perfil metabólico com outros parâmetros como a produção de etanol, o teor de taninos e o tratamento de desengorduramento fez-se uso das ferramentas disponíveis no software *MetaboAnalyst 6.0* (*Statistical Analysis metadata table*). Os metadados categorizados foram o tratamento (*in natura* ou desengordurado), o teor de

taninos (alto, intermediário e baixo), enquanto o teor de etanol foi um metadado contínuo. Na primeira análise procurou-se ter uma visão geral sobre a correlação dos metadados com o perfil metabólico. O programa propiciou um mapa de calor dos metadados bem como um mapa de calor da correlação. Por estes diagramas é possível identificar os metadados altamente dependentes entre si. Os diagramas mostraram que o teor de taninos se encontra altamente correlacionado ao teor de etanol, e que o tratamento de desengorduramento não está correlacionado. Ou seja, o teor de taninos tem maior influência na eficiência da produção de etanol que o desengorduramento, cuja correlação está apresentada na Figura 19.

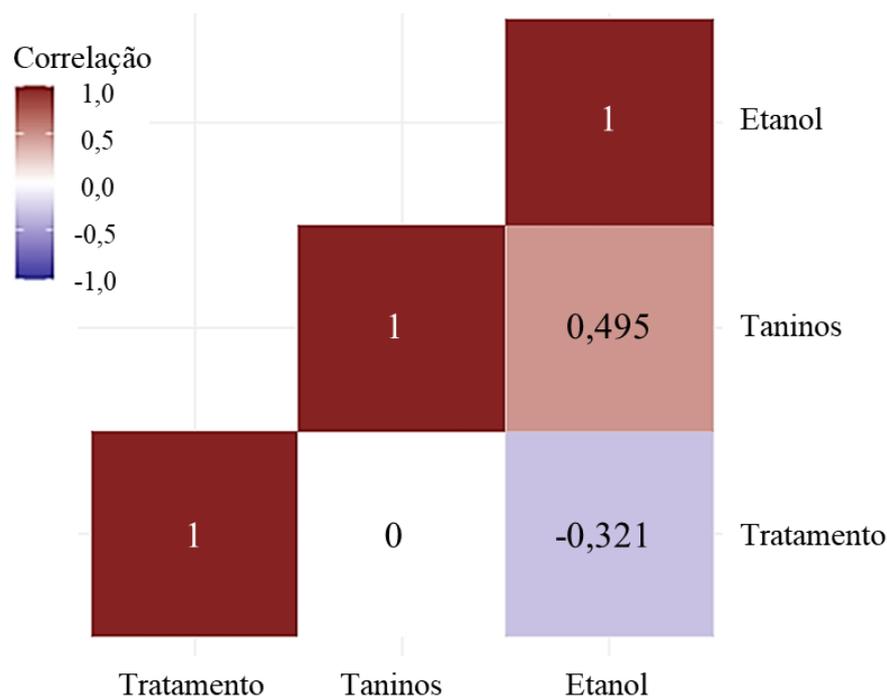


Figura 19. Diagrama das correlações entre os metadados tratamento (grãos in natura e desengordurados), teor de taninos (alto, intermediário e baixo) e teor de etanol.

Os dados podem ser submetidos a uma análise univariada com todas as correlações sendo analisadas concomitantemente ou obtendo-se a correlação parcial entre o metaboloma e os diferentes metadados de interesse. Ao fazer a correlação optou-se por utilizar o coeficiente de correlação de postos de Spearman, que é uma medida de correlação não-paramétrica. Ao contrário do coeficiente de correlação de Pearson, não requer a suposição que a relação entre as variáveis é linear, nem requer que as variáveis sejam quantitativas. A Figura 20 mostra os 25 *bins* mais fortemente correlacionados, sendo possível observar a dependência do teor de etanol ao teor de taninos, ou seja, a correlação parcial entre as variáveis (*bins*), o metadado de interesse (teor de etanol) e a covariável de interesse (teor de taninos).

Entre as variáveis de importância correlacionadas positivamente destacam-se duas em região de hidrogênios metílicos ou metilênicos ( $\delta$  1,66 e 1,94) entre outras que se situam na região de aromáticos ( $\delta$  7,3-7,82). Os *bins*  $\delta$  1,66 e 1,94 são intensos nas amostras *in natura* e quase não aparece nas amostras desengorduradas. Estes sinais são característicos de hidrogênios metilênicos de cadeias hidrocarbônicas como aquelas presentes em ácidos graxos ( $\delta$  1,66), bem como da presença de acetato ( $\delta$  1,94).

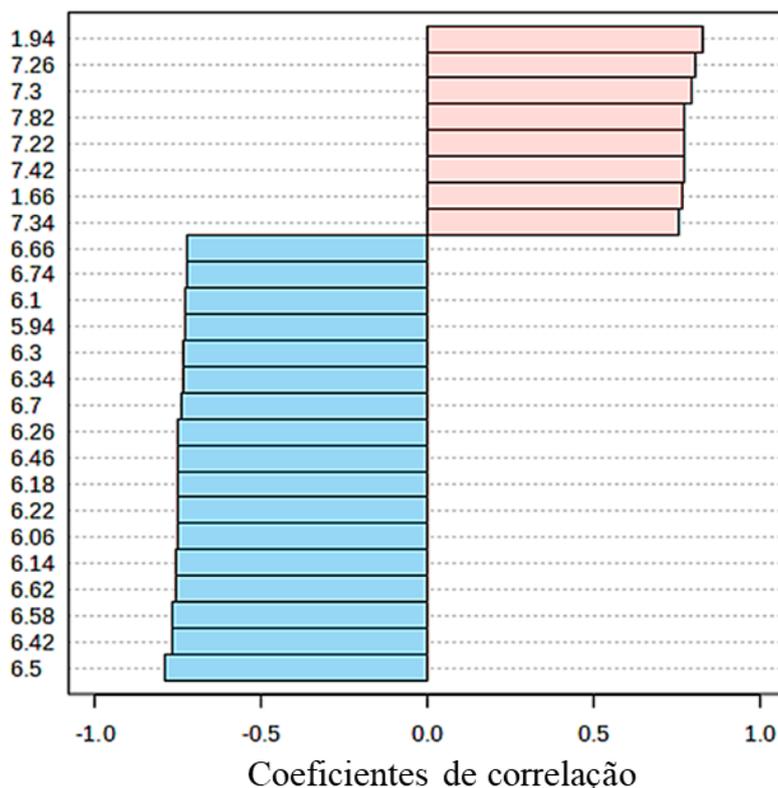


Figura 20. Os 25 principais *bins* correlacionados entre as variáveis (*bins*), os metadado de interesse (teor de etanol) e a covariável de interesse (teor de taninos). Em azul, deslocamentos químicos que afetam negativamente o teor de etanol. Em rosa, os que afetam positivamente.

Quando se analisa as variáveis correlacionadas negativamente ao teor de etanol (em azul, Figura 20) observa-se que estas se situam entre  $\delta$  5,94 e 6,74 e são condizentes com os sinais característicos de unidades flavan-3-ol presentes em proantocianidinas.

Ao se comparar os espectros de RMN de  $^1\text{H}$  de amostras *in natura* e desengorduradas detectamos os sinais (Mandrone et al., 2021; Nielsen et al., 2016) referentes aos ácidos graxos (Figura 25, apêndice B, página 83).

O híbrido BR501 foi o que apresentou a melhor eficiência para a produção de etanol, com 99,29% para amostras *in natura* e 98,72% para as amostras desengorduradas. O segundo genótipo com melhor eficiência para a produção de etanol foi a linhagem CMSXS180, com 92,31% para amostras *in natura* e 94,93% para as amostras desengorduradas. Os dois genótipos foram os que apresentaram menor teor em fenólicos, sendo o genótipo BR501 o que apresentou menor teor entre os dois. O espectro de RMN de  $^1\text{H}$  do BR501 foi comparado com o espectro do CMSXS180 (Figura 21) e foi possível identificar um menor número de sinais na região de hidrogênios aromáticos característicos de flavonoides. O único sinal que se destaca nessa região foi em  $\delta$  8,47 atribuído ao ácido fórmico. No espectro de RMN de  $^1\text{H}$  para a linhagem CMSXS180 foi possível observar alguns sinais condizentes com aqueles correlacionados positivamente à produção de etanol, a saber  $\delta$  7,81 (d,  $J = 8,2$  Hz), 7,39 (d,  $J = 8,7$  Hz), 7,29 (d,  $J = 8,6$  Hz), 7,10 (d,  $J = 8,7$  Hz), 6,84 (m), 6,53 (s) e 5,96 (s) (Figura 21).

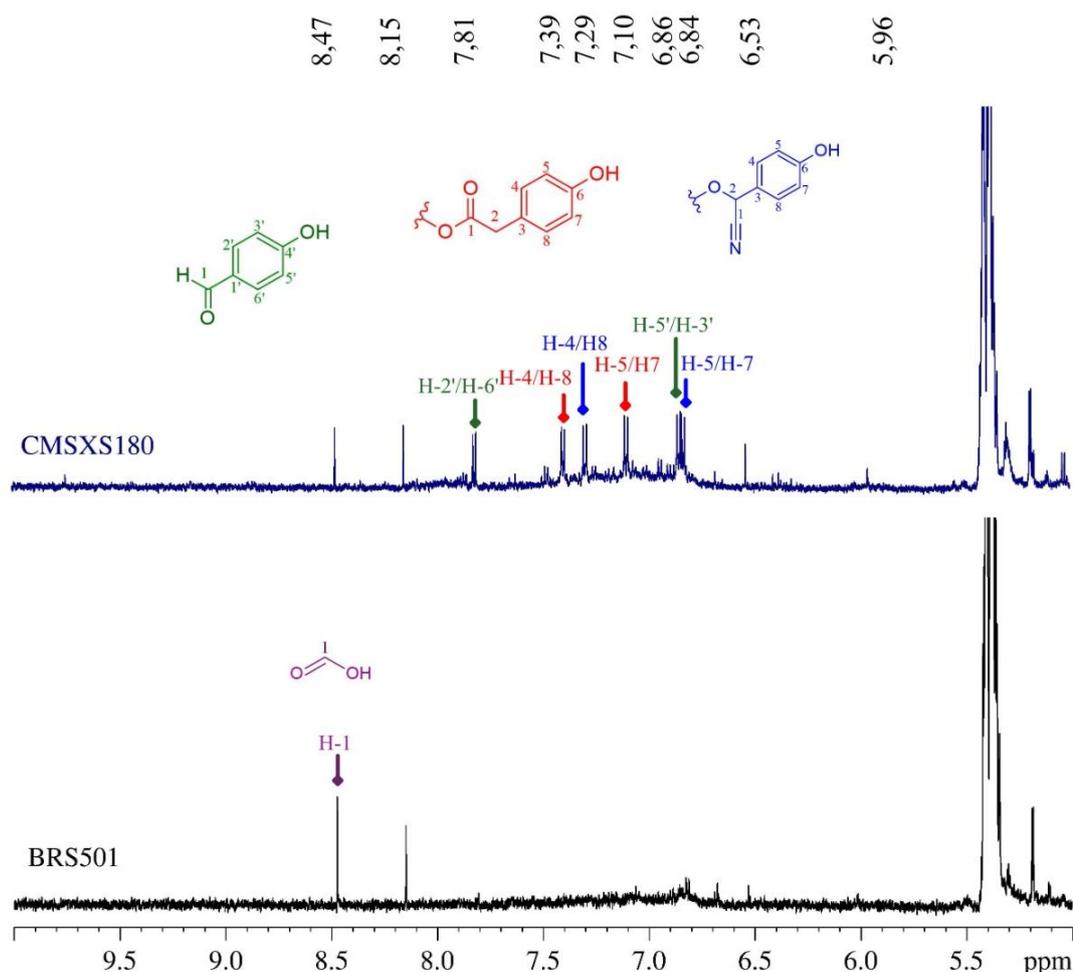


Figura 21. Espectro de RMN de  $^1\text{H}$  das amostras CMSXS180 (A) e BR501 (B) com expansão na região entre  $\delta$  5,0-10,0 (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0).

Pelo mapa de contornos COSY e pelo  $J$ -resolvido (Figura 22) para o genótipo de sorgo CMSX180 *in natura* é possível observar o acoplamento *orto* entre os sinais em  $\delta$  7,81 e 6,86; entre  $\delta$  7,39 e 7,10, entre  $\delta$  7,29 e 6,84. Somente acoplamentos *orto* não são observados para hidrogênios em derivados flavonoídicos, levando a proposição de outro tipo de composto aromático *para*-substituído. Ao se fazer uma busca na literatura sobre os metabólitos secundários presentes em sorgo, e se comparar os deslocamentos químicos dos sistemas aromáticos, detectou-se a possibilidade destes compostos serem derivados da tirosina, do ácido *p*-hidroxibenzoico, do *p*-hidroxibenzaldeído e do ácido *p*-hidroxifenilacético (Mandrone et al., 2021). Os sinais em  $\delta$  7,81 e 6,86 podem ser atribuídos ao sistema AA'BB' do *p*-hidroxibenzaldeído ou do ácido *p*-hidroxibenzoico, os sinais em  $\delta$  7,39 e 7,10 são condizentes com a presença do ácido *p*-glicosiloxy-2-hidroxifenilacético, e finalmente o sistema  $\delta$  7,29 e 6,84 poderia ser atribuído à durrina ou à *p*-hidroxymandelonitrila.

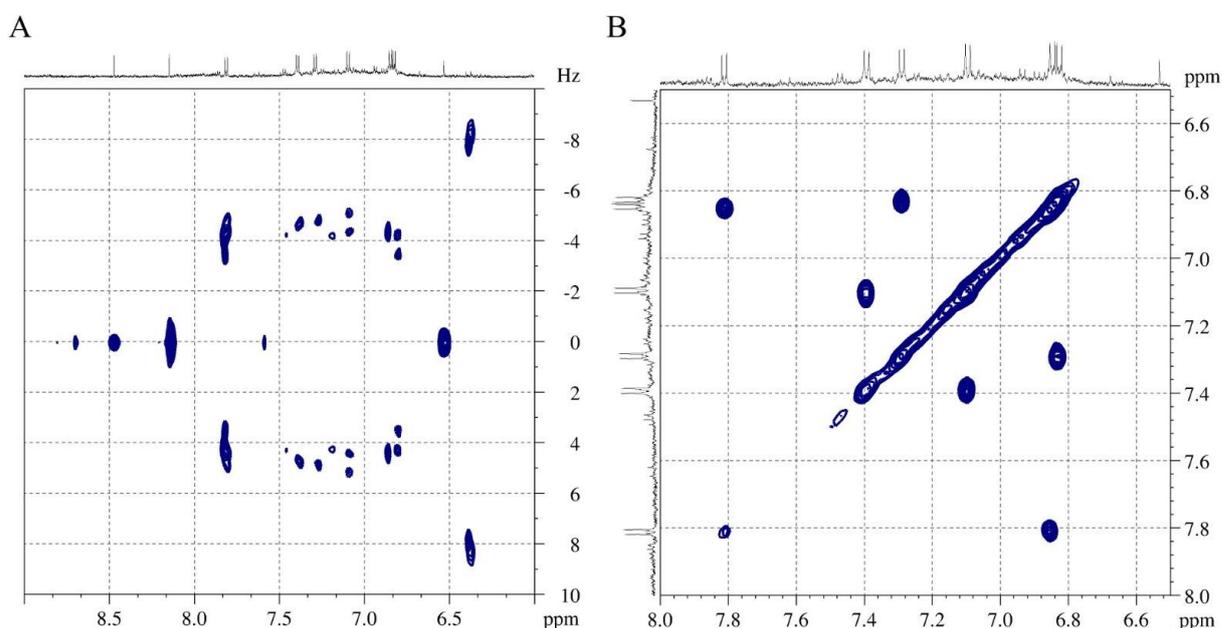


Figura 22. A) Mapa de contornos  $J$ -resolvido expansão em  $\delta$  6,0-8,6; B) Mapa de contornos COSY para o genótipo de sorgo CMSX180 *in natura* com a expansão entre  $\delta$  6,0 e 8,1 (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0).

Foi obtido um experimento de HSQC que possibilitou a elucidação dos carbonos, confirmando essas três estruturas (Figura 23).

Esses metabólitos já foram descritos para o sorgo anteriormente (Mandrone et al., 2021; Nielsen et al., 2016). Durrina é um heterosídeo cianogênico que é biossintetizada nos primeiros estágios de desenvolvimento do *S. bicolor* (L.) Moench como um mecanismo de defesa contra insetos e herbívoros generalistas, sendo ainda uma forma de estocar nitrogênio e uma forma de resistir a estresses abióticos, como a seca. Durante o desenvolvimento, a durrina é catabolizada e o ácido *p*-hidroxifenilacético seria o principal produto desse catabolismo, podendo também serem obtidos a *p*-hydroxymandelonitrila, o ácido *p*-hidroxibenzoico e o *p*-hidroxibenzaldeído.

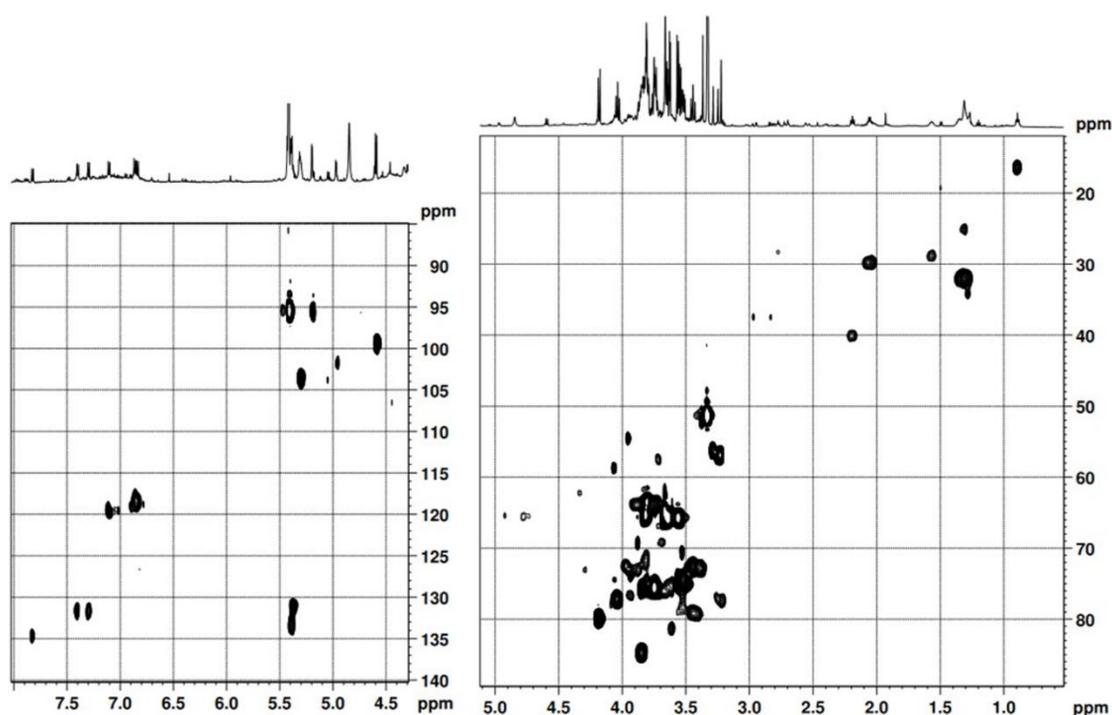


Figura 23. Mapa de contornos HSQC da cultivar CMSXS180 *in natura* destacando sinais atribuídos ao ácido *p*-glicosiloxy-2-hidroxifenilacético ( $\delta$  7,39 e 7,10), à durrina ( $\delta$  7,29 e 6,84) e ao ácido *p*-hidroxibenzoico ( $\delta$  7,81 e 6,86). (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0).

Quando se analisa os bins correlacionados negativamente ao teor de etanol, observou-se que o espectro de RMN de  $^1\text{H}$  da linhagem SC084 (Figura 26, apêndice B, página 84) apresentou todas as variáveis correlacionadas negativamente com o teor de etanol, sugerindo que estes compostos interferem de alguma forma no processo de fermentação.

Na tentativa de identificarmos as variáveis importantes relacionadas aos metadados, aplicamos o método de classificação supervisionado floresta aleatória, ou *random forest* (RF), incluído no módulo de análise estatística do *MetaboAnalyst* 6.0. O *random forest*, é um método de aprendizagem de máquina que pode ser usado para classificar e regredir variáveis binárias

ou contínuas. Este método é baseado em construir vários modelos de árvores de decisão, a partir de diferentes subconjuntos dos dados de treinamento e, em seguida, combinar os resultados destes modelos para melhorar as estatísticas de precisão dos resultados. As variáveis discriminatórias são ranqueadas no RF por sua contribuição para a precisão da classificação, isto é, através da diminuição média na precisão (MDA – *Mean Decrease Accuracy*) é possível estimar qual significância estatística daquela variável para a classificação de determinado grupo gerado de um grande conjunto de dados.

A matriz gerada a partir dos espectros de RMN aliada às informações dos metadados foram comparadas por seu desempenho discriminatório com mostrado na Figura 24 A e B. A precisão média da diminuição (MDA) é a medida do desempenho do modelo sem cada metabólito. Um valor mais alto indica a importância desse metabólito na previsão do grupo (teor de taninos *versus* teor de etanol). A remoção desse metabólito faz com que o modelo perca a precisão na previsão. Assim observa-se que os sinais característicos de unidades flavan-3-ol presentes em proantocianidinas são importantes na classificação das amostras com alto teor de taninos e, aliado aos resultados da análise anterior, sugerem uma alta previsão para o modelo onde a detecção destes deslocamentos químicos indicariam uma menor eficiência na produção de etanol.

Estes dados precisam ser melhor explorados, mas são um indício que a metabolômica por RMN pode fornecer informações interessantes sobre os cultivares e propiciar modelos para a previsão de uma maior ou menor eficiência na produção de etanol.

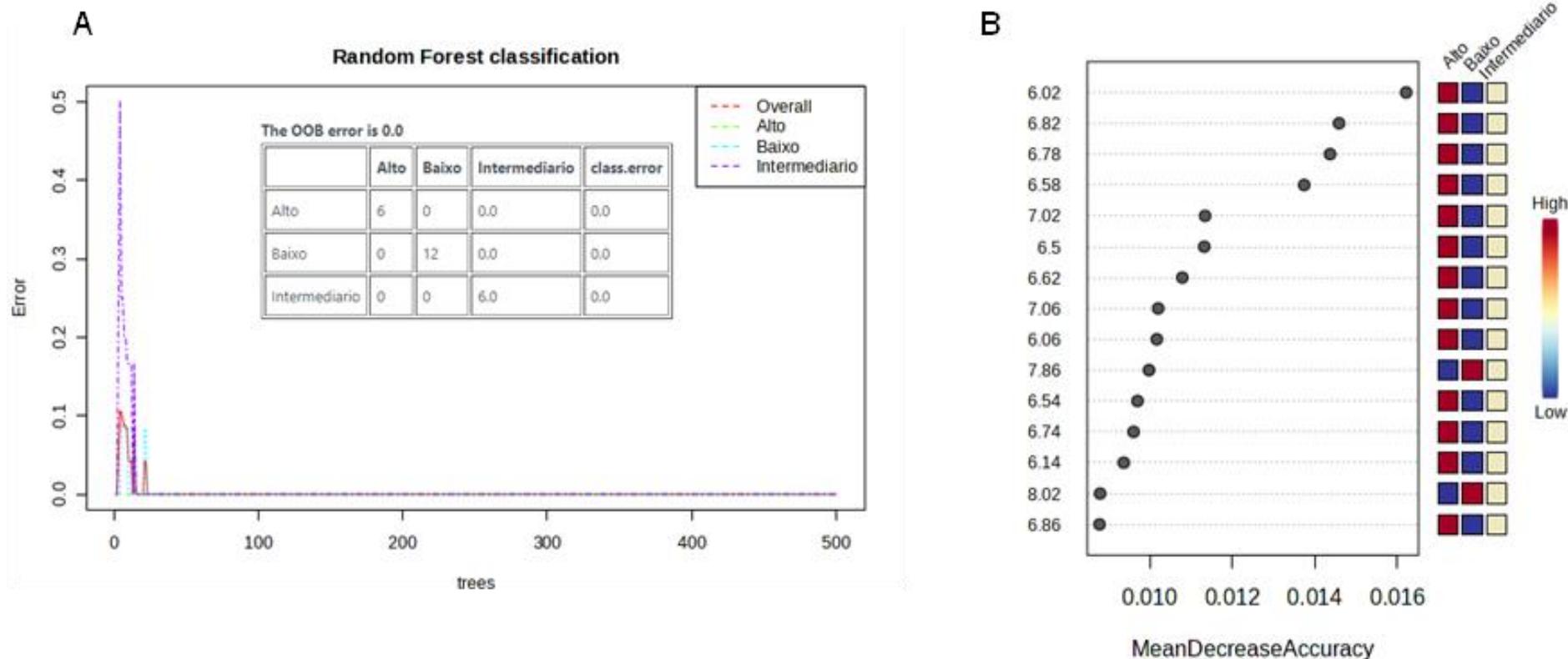


Figura 24. A) Modelo de classificação *random forest* (RF) mostrando o erro medido para cada classe usando o algoritmo de aprendizado de máquina, B) Bins classificados por suas contribuições para a precisão da classificação (diminuição média na precisão).

## 6. CONCLUSÃO

Quatro fenótipos de *Sorghum bicolor* L. Moench, foram empregados na produção de etanol, com rendimentos variando de 83,8 a 99,3 % para as amostras com maior e menor teor de compostos fenólicos, respectivamente. A remoção de lipídeos através do tratamento de desengorduramento dos grãos resultou em um aumento percentual médio de 4,9 % no teor de amido total nas amostras, porém esse aumento não refletiu numa maior eficiência na produção de etanol, uma vez que no rendimento final as amostras desengorduradas, na média, apresentaram uma queda no rendimento do processo.

Avaliando-se as influências da composição (teor de taninos) e tratamento (amostras *in natura* ou desengorduradas) foi possível determinar que o teor de taninos tem maior influência na eficiência da produção de etanol do que o processo de desengorduramento, podendo explicar a ausência de um padrão de aumento ou diminuição no rendimento da fermentação para as amostras desengorduradas.

A metabolômica por RMN se mostrou uma ferramenta eficiente na diferenciação do perfil metabólico dos genótipos de sorgo. Ela forneceu informações que possibilitaram identificar em SC084, previamente reportada como sendo uma cultivar híbrida com taninos, metabólitos diferentes dos demais, apresentando principalmente compostos com deslocamentos químicos na região de flavonoides. Foi possível identificar também um perfil diferenciado para a linhagem do genótipo CMSXS180, linhagem reportada com ausência de taninos, mas que apresentou sinais condizentes com a presença da durrina e produtos de seu catabolismo. Os espectros de RMN de  $^1\text{H}$  do híbrido BR501 revelaram que o processo de melhoramento realmente possibilitou a remoção de vias metabólicas decorrentes da biossíntese envolvendo a tirosina.

A análise quimiométrica revelou que a presença da durrina e metabólitos de seu catabolismo não interfere na produção de bioetanol e coloca, desta forma, a possibilidade de aproveitamento de culturas de sorgo com esses metabólitos possam ser direcionadas para a obtenção de biocombustível.

Em resumo, conclui-se que a presença de compostos fenólicos, em especial taninos, é um fator determinante no rendimento do processo de obtenção do etanol, podendo estar

interferindo no processo de hidrólise, sacarificação ou fermentação e diminuindo o rendimento da reação.

Foi possível perceber, por este estudo, que a espectroscopia de RMN é adequada na diferenciação metabólica das cultivares de sorgo e que a técnica, associada a ferramentas quimiométricas, possui potencial de aplicação em estudos metabolômicos visando a associação de perfis metabólicos à eficiência do processo de produção de etanol. A RMN de  $^1\text{H}$  forneceu informações consistentes e confiáveis sobre o conteúdo de fenólicos e outros compostos importantes do perfil metabólico do sorgo de forma mais rápida e eficiente que os métodos espectroscópicos usualmente utilizados. Este estudo representa um ponto de partida para a utilização desta ferramenta na cadeia produtiva do sorgo visando sua utilização como biomassa na matriz energética.

## REFERÊNCIAS

Angelo, P. M., & Jorge, N. (2007). Compostos fenólicos em alimentos - Uma breve revisão. *Rev. Inst. Adolfo Lutz*, 66(1), 1–9.

AOCS. (2005). Official Method AM5-04. Rapid determination of oil/fat utilizing high-temperature solvent extraction. *Official Methods and Recommended Practices of the American Oil Chemists' Society*.

Awika, J. M., McDonough, C. M., & Rooney, L. W. (2005). Decorticating Sorghum To Concentrate Healthy Phytochemicals. *Journal of Agricultural and Food Chemistry*, 53(16), 6230–6234. <https://doi.org/10.1021/jf0510384>

Awika, J. M., Rooney, L. W., & Waniska, R. D. (2004). Properties of 3-Deoxyanthocyanins from Sorghum. *Journal of Agricultural and Food Chemistry*, 52(14), 4388–4394. <https://doi.org/10.1021/jf049653f>

Awika, J. M., Rooney, L. W., Wu, X., Prior, R. L., & Cisneros-Zevallos, L. (2003). Screening Methods To Measure Antioxidant Activity of Sorghum (*Sorghum bicolor*) and Sorghum Products. *Journal of Agricultural and Food Chemistry*, 51(23), 6657–6662. <https://doi.org/10.1021/jf034790i>

Barison, A., Pereira da Silva, C. W., Campos, F. R., Simonelli, F., Lenz, C. A., & Ferreira, A. G. (2010). A simple methodology for the determination of fatty acid composition in edible oils through <sup>1</sup>H NMR spectroscopy. *Magnetic Resonance in Chemistry*, 48(8), 642–650. <https://doi.org/10.1002/mrc.2629>

Bhering, L. L. (2017). Rbio: A tool for biometric and statistical analysis using the R platform. *Crop Breeding and Applied Biotechnology*, 17(2), 187–190. <https://doi.org/10.1590/1984-70332017v17n2s29>

BNDES, & CGEE. (2008). *Bioetanol de cana-de-açúcar: energia para o desenvolvimento sustentável* (1 ed). BNDES.

Brereton, R. G. (2003). *Chemometrics: data analysis for the laboratory and chemical plant*. John Wiley & Sons.

Canuto, G., Costa, J. L., Cruz, P., Souza, A., Faccio, A., Klassen, A., Rodrigues, K., & Tavares, M. (2018). Metabolômica: definições, estado-da-arte e aplicações representativas. *Química Nova*, 41(1), 75–91. <https://doi.org/10.21577/0100-4042.20170134>

Cardoso, W. S., Paes, M. C. D., Dias, S. R., Camporezi, J. D., Modolo, M. U., & Pinheiro, F. de A. (2020). *Metodologia científica: determinação do rendimento de produção de etanol de grãos de milho por via seca em escala laboratorial*.

- Cevallos-Cevallos, J. M., Reyes-De-Corcuera, J. I., Etxeberria, E., Danyluk, M. D., & Rodrick, G. E. (2009). Metabolomic analysis in food science: a review. *Trends in Food Science & Technology*, 20(11–12), 557–566. <https://doi.org/10.1016/j.tifs.2009.07.002>
- Chao, C., Liang, S., Zhang, Z., Gidley, M. J., Liu, Y., & Wang, S. (2024). New Insight into the Effects of Endogenous Protein and Lipids on the Enzymatic Digestion of Starch in Sorghum Flour. *Foods*, 13(5), 663. <https://doi.org/10.3390/foods13050663>
- Choi, H.-K., Choi, Y. H., Verberne, M., Lefeber, A. W. M., Erkelens, C., & Verpoorte, R. (2004). Metabolic fingerprinting of wild type and transgenic tobacco plants by 1H NMR and multivariate analysis technique. *Phytochemistry*, 65(7), 857–864. <https://doi.org/10.1016/j.phytochem.2004.01.019>
- Chuck-Hernandez, C., Peralta-Contreras, M., Bando-Carranza, G., Vera-Garcia, M., Gaxiola-Cuevas, N., Tamayo-Limon, R., Cardenas-Torres, F., Perez-Carrillo, E., & Serna-Saldivar, S. O. (2012). Bioconversion into ethanol of decorticated red sorghum (*Sorghum bicolor* L. Moench) supplemented with its phenolic extract or spent bran. *Biotechnology Letters*, 34(1), 97–102. <https://doi.org/10.1007/s10529-011-0752-z>
- CONAB (2023). Boletim da safra de grãos/produção e balanço de oferta e demanda de grãos. Disponível em: <https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos>. Acessado em: 15 de julho de 2024.
- Cunniff, P., & Washington, D. (1995). *Official methods of analysis of AOAC International: Vol. v.1* (16. ed.). AOAC International.
- Dumont, D., Danielato, G., Chastellier, A., Hibrand Saint Oyant, L., Fanciullino, A.-L., & Lukan, R. (2020). Multi-Targeted Metabolic Profiling of Carotenoids, Phenolic Compounds and Primary Metabolites in Goji (*Lycium* spp.) Berry and Tomato (*Solanum lycopersicum*) Reveals Inter and Intra Genus Biomarkers. *Metabolites*, 10(10), 422. <https://doi.org/10.3390/metabo10100422>
- Empresa de Pesquisa Energética. (2023). *Análise de Conjuntura dos Biocombustíveis: Ano 2022*. <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/analise-de-conjuntura-dos-biocombustiveis-2022>
- Emwas, A.-H., Saccenti, E., Gao, X., McKay, R. T., dos Santos, V. A. P. M., Roy, R., & Wishart, D. S. (2018). Recommended strategies for spectral processing and post-processing of 1D 1H-NMR data of biofluids with a particular focus on urine. *Metabolomics*, 14(3), 31. <https://doi.org/10.1007/s11306-018-1321-4>
- Ferreira, M. M. C. (2015). *Quimiometria: conceitos, métodos e aplicações*. Editora Unicamp.
- Foglia, F., Drapcho, C., & Nghiem, J. (2022). Effectiveness of Tannin Removal by Alkaline Pretreatment on Sorghum Ethanol Production. *Fermentation*, 8(6), 274. <https://doi.org/10.3390/fermentation8060274>

Gomes, P. C., & Simeone, M. L. F. (2012). *Comunicado Técnico 202: Determinação rápida de extrato etéreo utilizando extrator a alta temperatura.*

Guitton, Y., Tremblay-Franco, M., Le Corguillé, G., Martin, J.-F., Pétéra, M., Roger-Mele, P., Delabrière, A., Goulitquer, S., Monsoor, M., Duperier, C., Canlet, C., Servien, R., Tardivel, P., Caron, C., Giacomoni, F., & Thévenot, E. A. (2017). Create, run, share, publish, and reference your LC–MS, FIA–MS, GC–MS, and NMR data analysis workflows with the Workflow4Metabolomics 3.0 Galaxy online infrastructure for metabolomics. *The International Journal of Biochemistry & Cell Biology*, *93*, 89–101. <https://doi.org/10.1016/j.biocel.2017.07.002>

Gupta, N., Yadav, D. K., Gautam, S., Kumar, A., Kumar, D., & Prasad, N. (2023). Nuclear Magnetic Resonance-Based Metabolomics Approach Revealed the Intervention Effect of Using Complementary and Alternative Medicine (CAM) by CKD Patients. *ACS Omega*, *8*(8), 7722–7737. <https://doi.org/10.1021/acsomega.2c06469>

Hausman, C. (2012). Biofuels and Land Use Change: Sugarcane and Soybean Acreage Response in Brazil. *Environmental and Resource Economics*, *51*(2), 163–187. <https://doi.org/10.1007/s10640-011-9493-7>

Jaumot, J., Bedia, C., & Tauler, R. (2018). *Data Analysis for Omic Sciences: Methods and Applications: Vol. v. 82* (1 ed). Elsevier.

Jesus, M. dos P. M. (2018). *Purificação e elucidação estrutural de acetogeninas isoladas de Annona mucosa (Jacq.) através de RMN.* Universidade Federal de Alagoas.

Kaiser, C. R. (2000). RMN 2D: Detecção inversa e gradiente de campo na determinação estrutural de compostos orgânicos. *Química Nova*, *23*(2), 231–236. <https://doi.org/10.1590/S0100-40422000000200014>

Kim, H. K., Choi, Y. H., & Verpoorte, R. (2010). NMR-based metabolomic analysis of plants. *Nature Protocols*, *5*(3), 536–549. <https://doi.org/10.1038/nprot.2009.237>

Lacerda, M. S., Junior, J. C. A. S., & Emiliano, P. C. (2022). Expansão da cana-de-açúcar: uma investigação via cópulas. , *31*(2), p.33. In *Revista de Política Agrícola* (pp. 33–50). Embrapa.

Last, R. L., Jones, A. D., & Shachar-Hill, Y. (2007). Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology*, *8*(2), 167–174. <https://doi.org/10.1038/nrm2098>

Lindon, J. C., & Nicholson, J. K. (2008). Spectroscopic and Statistical Techniques for Information Recovery in Metabonomics and Metabolomics. *Annual Review of Analytical Chemistry*, *1*(1), 45–69. <https://doi.org/10.1146/annurev.anchem.1.031207.113026>

Ludwig, C., & Viant, M. R. (2010). Two-dimensional *J*-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochemical Analysis*, *21*(1), 22–32. <https://doi.org/10.1002/pca.1186>

- Macedo, A. A., Medeiros, R. G., Silvério, T. A. B., Nelson, D. L., Souza Oliveira, D. C., & dos Reis, A. B. (2020). Possibilities and Prospects Regarding Ethanol Production from Saccharin Sorghum [*Sorghum bicolor* (L.) Moench]. *SN Applied Sciences*, 2(12), 2186. <https://doi.org/10.1007/s42452-020-03912-7>
- Machado, A. R. T. (2014). *Metabolômica por Ressonância Magnética Nuclear de Annona spp para o Controle de Nematoides*. Universidade Federal de Minas Gerais.
- Mandrone, M., Chiochio, I., Barbanti, L., Tomasi, P., Tacchini, M., & Poli, F. (2021). Metabolomic Study of Sorghum (*Sorghum bicolor*) to Interpret Plant Behavior under Variable Field Conditions in View of Smart Agriculture Applications. *Journal of Agricultural and Food Chemistry*, 69(3), 1132–1145. <https://doi.org/10.1021/acs.jafc.0c06533>
- Manochio, C., Andrade, B. R., Rodriguez, R. P., & Moraes, B. S. (2017). Ethanol from biomass: A comparative overview. *Renewable and Sustainable Energy Reviews*, 80, 743–755. <https://doi.org/10.1016/j.rser.2017.05.063>
- Maraschin, M., Zeggio, A. R. S., Tomazzoli, M. M., Oliveira, S. K., Ramlov, F., Veleirinho, M. B., & Rocha, M. (2017). *Metabolômica e quimiometria como ferramentas para análises químico (bio) diversas*.
- Marchev, A. S., Vasileva, L. V., Amirova, K. M., Savova, M. S., Balcheva-Sivenova, Z. P., & Georgiev, M. I. (2021). Metabolomics and health: from nutritional crops and plant-based pharmaceuticals to profiling of human biofluids. *Cellular and Molecular Life Sciences*, 78(19–20), 6487–6503. <https://doi.org/10.1007/s00018-021-03918-3>
- Moco, S. (2022). Studying Metabolism by NMR-Based Metabolomics. *Frontiers in Molecular Biosciences*, 9. <https://doi.org/10.3389/fmolb.2022.882487>
- Monteiro, J. M., Albuquerque, U. P. de, Araújo, E. de L., & Amorim, E. L. C. de. (2005). Taninos: uma abordagem da química à ecologia. *Química Nova*, 28(5), 892–896. <https://doi.org/10.1590/S0100-40422005000500029>
- Nielsen, L. J., Stuart, P., Pičmanová, M., Rasmussen, S., Olsen, C. E., Harholt, J., Møller, B. L., & Bjarnholt, N. (2016). Dhurrin metabolism in the developing grain of *Sorghum bicolor* (L.) Moench investigated by metabolite profiling and novel clustering analyses of time-resolved transcriptomic data. *BMC Genomics*, 17(1), 1021. <https://doi.org/10.1186/s12864-016-3360-4>
- Nitsch, M. (1991). O programa de biocombustíveis Proálcool no contexto da estratégia energética brasileira. *Brazilian Journal of Political Economy*, 11(2), 274–299. <https://doi.org/10.1590/0101-31571991-0712>
- ONU. (2016). *Sobre o nosso trabalho para alcançar os Objetivos de Desenvolvimento Sustentável no Brasil*. Sobre o nosso trabalho para alcançar os Objetivos de Desenvolvimento Sustentável no Brasil

- Palacios, C. E., Nagai, A., Torres, P., Rodrigues, J. A., & Salatino, A. (2021). Contents of tannins of cultivars of sorghum cultivated in Brazil, as determined by four quantification methods. *Food Chemistry*, *337*, 127970. <https://doi.org/10.1016/j.foodchem.2020.127970>
- Panero, F. S., Vieira, M. F. P., Cruz, Â. M. F., Moura, M. F. V., & Silva, H. E. B. da. (2009). Aplicação da análise exploratória de dados na discriminação geográfica do quiabo do Rio Grande do Norte e Pernambuco. *Eclética Química*, *34*(3), 33–40. <https://doi.org/10.1590/S0100-46702009000300004>
- Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., & Xia, J. (2022). Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nature Protocols*, *17*(8), 1735–1761. <https://doi.org/10.1038/s41596-022-00710-w>
- Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology*, *13*(4), 263–269. <https://doi.org/10.1038/nrm3314>
- Pavia, D. L., Lampman, G. M., Kriz, G. S., & Vyvyan, J. R. (2010). *Introdução à espectroscopia* (4 ed). Cengage Learning.
- Pereira Filho, & Rodrigues. (2015). *Sorgo: o produtor pergunta, a Embrapa responde* (1 ed). Embrapa .
- Pérez, M., Dominguez-López, I., & Lamuela-Raventós, R. M. (2023). The Chemistry Behind the Folin–Ciocalteu Method for the Estimation of (Poly)phenol Content in Food: Total Phenolic Intake in a Mediterranean Dietary Pattern. *Journal of Agricultural and Food Chemistry*, *71*(46), 17543–17553. <https://doi.org/10.1021/acs.jafc.3c04022>
- Pilon, A., Selegato, D., Fernandes, R., Bueno, P., Pinho, D., Carnevale Neto, F., Freire, R., Castro-Gamboa, I., Bolzani, V., & Lopes, N. (2020). Metabolômica de plantas: métodos e desafios. *Química Nova*. <https://doi.org/10.21577/0100-4042.20170499>
- Pilon, A., Vieira, N., Amaral, J., Monteiro, A., Silva, R., Spíndola, L., Castro-Gamboa, I., & Lopes, N. (2021). Redes moleculares: uma análise sobre anotações e descoberta de novos ativos. *Química Nova*. <https://doi.org/10.21577/0100-4042.20170777>
- Pimenta, L. P. S., Kim, H. K., Verpoorte, R., & Choi, Y. H. (2013). NMR-Based Metabolomics: A Probe to Utilize Biodiversity. In *Metabolomics Tools for Natural Product Discovery: Vol. v. 1055* (pp. 117–127). [https://doi.org/10.1007/978-1-62703-577-4\\_9](https://doi.org/10.1007/978-1-62703-577-4_9)
- Pinaki, D., Lhakpa, W., & Joginder, S. (2015). Simultaneous Saccharification and Fermentation (SSF), An Efficient Process for Bio-Ethanol Production: An overview. *Biosciences, Biotechnology Research Asia*, *12*(1), 87–100.

- Price, M. L., Van Scoyoc, S., & Butler, L. G. (1978). A critical evaluation of the vanillin reaction as an assay for tannin in sorghum grain. *Journal of Agricultural and Food Chemistry*, 26(5), 1214–1218. <https://doi.org/10.1021/jf60219a031>
- Ramchandran, D., Johnston, D. B., Tumbleson, M. E., Rausch, K. D., & Singh, V. (2015). Seasonal variability in ethanol concentrations from a dry grind fermentation operation associated with incoming corn variability. *Industrial Crops and Products*, 67, 155–160. <https://doi.org/10.1016/j.indcrop.2015.01.029>
- Rao, S., Santhakumar, A. B., Chinkwo, K. A., Wu, G., Johnson, S. K., & Blanchard, C. L. (2018). Characterization of phenolic compounds and antioxidant activity in sorghum grains. *Journal of Cereal Science*, 84, 103–111. <https://doi.org/10.1016/j.jcs.2018.07.013>
- RFA. (2024). *Ethanol Industry Outlook*. <https://ethanolrfa.org/resources/annual-industry-outlook>
- Rooney, L. W., & Miller, F. R. (1981). Variation in the structure and kernel characteristics of sorghum. In *Proceedings of the international symposium on sorghum grain quality* (Vol. 28, pp. 143–162).
- Sánchez, Ó. J., & Cardona, C. A. (2008). Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresource Technology*, 99(13), 5270–5295. <https://doi.org/10.1016/j.biortech.2007.11.013>
- Schaffert, R. E., & Rodrigues, J. A. S. (2014). *Fluxo gênico em sorgo*. Embrapa.
- Schofield, P., Mbugua, D. M., & Pell, A. N. (2001). Analysis of condensed tannins: a review. *Animal Feed Science and Technology*, 91(1–2), 21–40. [https://doi.org/10.1016/S0377-8401\(01\)00228-0](https://doi.org/10.1016/S0377-8401(01)00228-0)
- Shiao, T., & Shiao, M. (1989). Determination of fatty acid compositions of triacylglycerols by high resolution NMR spectroscopy. *Bot Bull Academia Sinica*, 30, 191–199.
- Sumner, L. W., Mendes, P., & Dixon, R. A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry*, 62(6), 817–836. [https://doi.org/10.1016/S0031-9422\(02\)00708-2](https://doi.org/10.1016/S0031-9422(02)00708-2)
- Taylor, J. R. N., Schober, T. J., & Bean, S. R. (2006). Novel food and non-food uses for sorghum and millets. *Journal of Cereal Science*, 44(3), 252–271. <https://doi.org/10.1016/j.jcs.2006.06.009>
- Trygg, J., Holmes, E., & Lundstedt, T. (2007). Chemometrics in Metabonomics. *Journal of Proteome Research*, 6(2), 469–479. <https://doi.org/10.1021/pr060594q>
- Valdes, C., Hjort, K., & Seeley, R. (2016). Brazil's Agricultural Land Use and Trade: Effects of Changes in Oil Prices and Ethanol Demand. *AgrEcon Search - Search in Agricultural & Applied Economics*.

- Van der Maarel, M. J. E. C., van der Veen, B., Uitdehaag, J. C. M., Leemhuis, H., & Dijkhuizen, L. (2002). Properties and applications of starch-converting enzymes of the  $\alpha$ -amylase family. *Journal of Biotechnology*, *94*(2), 137–155.  
[https://doi.org/10.1016/S0168-1656\(01\)00407-2](https://doi.org/10.1016/S0168-1656(01)00407-2)
- Wishart, D. S. (2008). Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology*, *19*(9), 482–493.  
<https://doi.org/10.1016/j.tifs.2008.03.003>
- Xia, J., Psychogios, N., Young, N., & Wishart, D. S. (2009). MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research*, *37*(Web Server), W652–W660. <https://doi.org/10.1093/nar/gkp356>
- Xia, J., & Wishart, D. S. (2011). Metabolomic Data Processing, Analysis, and Interpretation Using MetaboAnalyst. *Current Protocols in Bioinformatics*, *34*(1).  
<https://doi.org/10.1002/0471250953.bi1410s34>
- Zhou, Y., Lv, J., Yu, Z., Wang, Z., Li, Y., Li, M., Deng, Z., Xu, Q., Cui, F., & Zhou, W. (2022). Integrated metabolomics and transcriptomic analysis of the flavonoid regulatory networks in Sorghum bicolor seeds. *BMC Genomics*, *23*(1), 619.  
<https://doi.org/10.1186/s12864-022-08852-7>

## APÊNDICE A – Tabelas

Tabela 9. Teor de compostos fenólicos totais

| Amostra  | Embrapa<br>Teor de fenólicos totais<br>(mg GAE /g amostra) | CerQBio<br>Teor de fenólicos totais<br>(mg GAE /g amostra) |
|----------|--|--|
| SC084    | 37.26 ± 4,50   | 20,19 ± 0,83   |
| BRS305   | 22.15 ± 1,40   | 8,19 ± 0,60  |
| CMSXS180 | <LQ <sup>c</sup>   | 3,40 ± 0,03  |
| BR501    | <LQ <sup>c</sup>   | 2,92 ± 0,11  |

GAE (Ácido Gálico Equivalente), LQ (Limite de Quantificação) = 0,178 mg GAE/g amostra.

Tabela 10. Teores de glicose, glicerol e etanol para cada grama de amostra de sorgo *in natura* e desengorduradas após o processo de fermentação

| Amostra    | Glicose<br>(mg/g) | Glicerol<br>(mg/g) | Etanol<br>(mg/g) | Etanol<br>(mL/g) |
|------------|-------------------|--------------------|------------------|------------------|
| SC084      | 1.00 ± 0.31       | 31.31 ± 0.97       | 277.21 ± 3.94    | 0.351 ± 0,005    |
| SC084_d    | 0.77 ± 0.15       | 30.38 ± 1.07       | 284.50 ± 8.93    | 0.361 ± 0,011    |
| BRS305     | 1.97 ± 2.33       | 30.98 ± 0.95       | 304.31 ± 4.81    | 0.386 ± 0,006    |
| BRS305_d   | 1.34 ± 0.74       | 32.45 ± 0.24       | 314.18 ± 10.19   | 0.398 ± 0,013    |
| CMSXS180   | 1.06 ± 0.22       | 32.14 ± 0.91       | 313.94 ± 5.16    | 0.398 ± 0,007    |
| CMSXS180_d | 1.84 ± 0.33       | 32.95 ± 0.66       | 331.86 ± 5.72    | 0.421 ± 0,007    |
| BR501      | 1.15 ± 0.22       | 33.68 ± 0.91       | 317.94 ± 6.04    | 0.403 ± 0,008    |
| BR501_d    | 1.56 ± 0.43       | 35.28 ± 1.03       | 333.64 ± 13.59   | 0.423 ± 0,017    |

## APÊNDICE B – Espectros de RMN

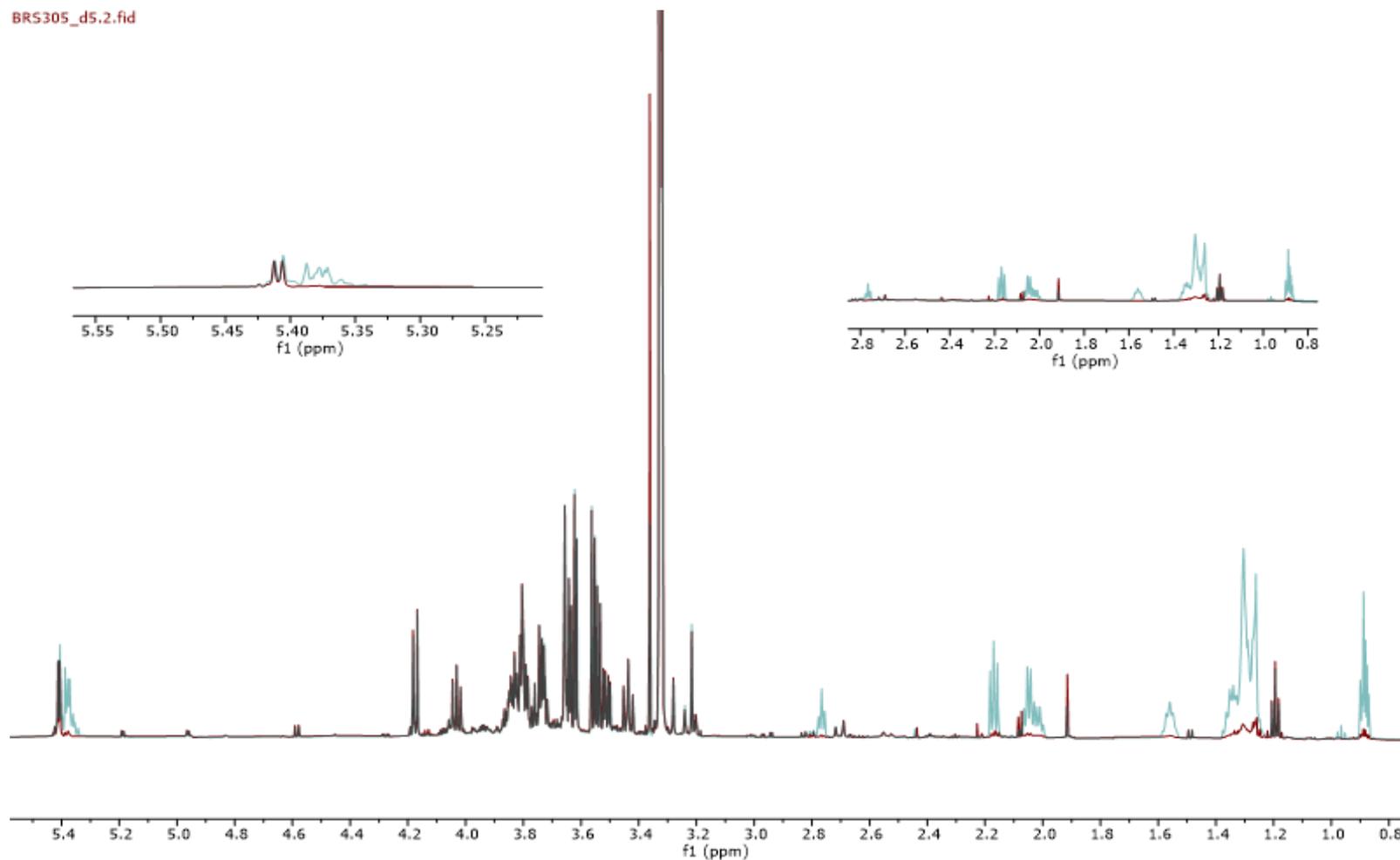


Figura 25. Espectros de RMN de <sup>1</sup>H da amostra de sorgo BRS305 *in natura* (verde) e desengordurada (marrom). (600 MHz, Metanol-*d*<sub>4</sub> e tampão KH<sub>2</sub>PO<sub>4</sub> em D<sub>2</sub>O com 0,01% TSP-*d*<sub>4</sub> (m/m), pH 6,0).

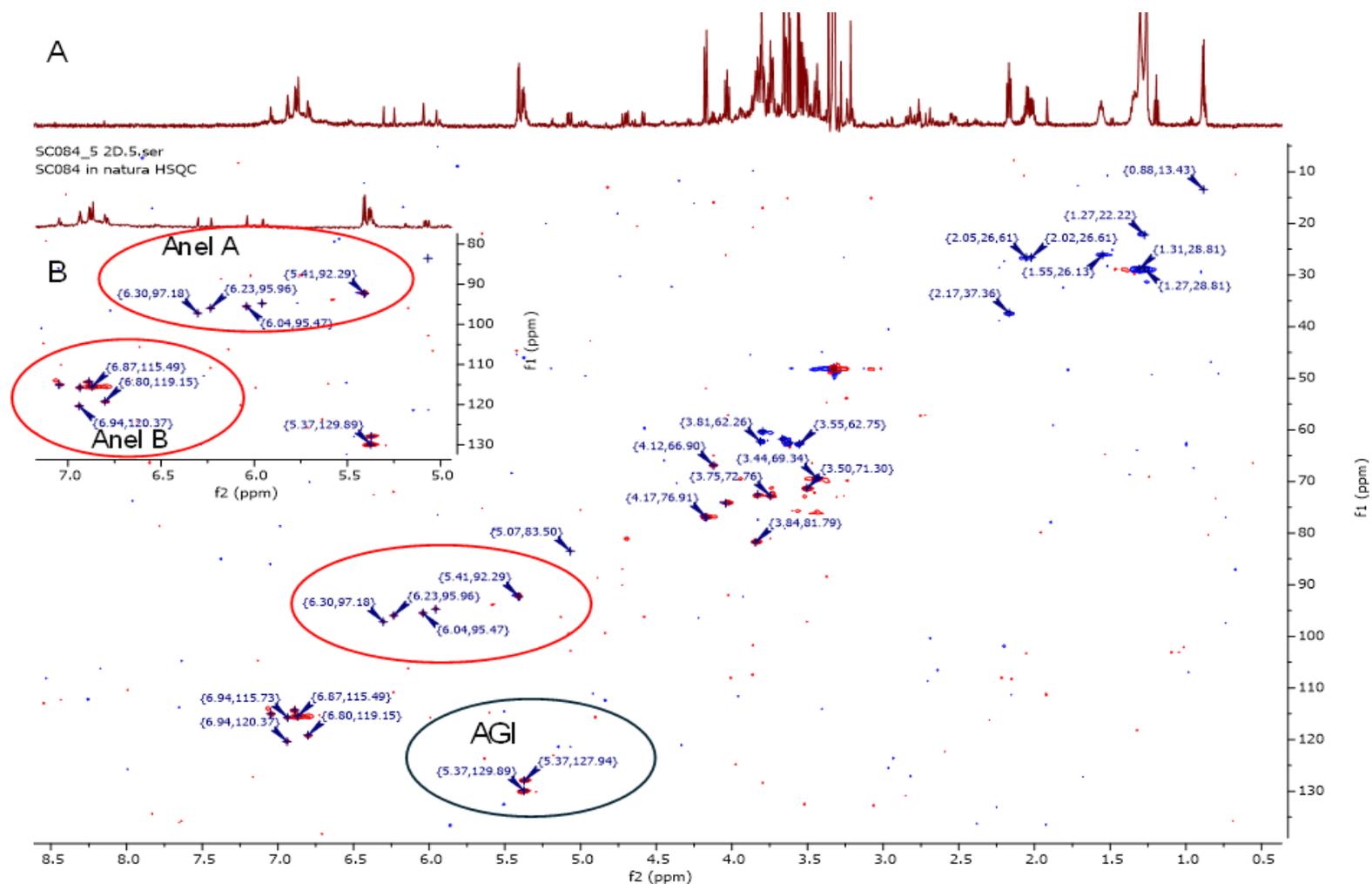
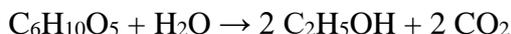


Figura 26. A) Mapa de contornos HSQC da cultivar SC084 *in natura* destacando a região do anel A de unidades flavonoídicas e sinais característicos de ácidos graxos insaturados (AGI); B) expansão entre  $\delta$  5 e 7,5 destacando a região de sinais dos anéis A e B de unidades flavonoídicas. (600 MHz, Metanol- $d_4$  e tampão  $\text{KH}_2\text{PO}_4$  em  $\text{D}_2\text{O}$  com 0,01% TSP- $d_4$  (m/m), pH 6,0).

### APÊNDICE C – Cálculos

- Passo 1: determinar o rendimento teórico da reação



1 mol de amido produz 2 mol de etanol, ou seja, 162 g de amido produz 92 g

Pela lei da conservação de massas, temos que o número de mol do reagente deve se equivaler ao dos produtos,  $n = n$ . Aplicando o conceito onde o número de mol é  $m/\text{MM}$ , onde  $m$  = massa e  $\text{MM}$  = massa molar do átomo, temos que:

$$\frac{m_{\text{etanol, g}}}{\text{MM}_{\text{etanol}}} = \frac{m_{\text{amido teórico, g}}}{\text{MM}_{\text{amido}}}$$

$$m_{\text{etanol, g}} = \frac{m_{\text{amido teórico, g}}}{\text{MM}_{\text{amido}}} \times \text{MM}_{\text{etanol}}$$

$$m_{\text{etanol, g}} = \frac{m_{\text{amido teórico, g}}}{162 \text{ g}} \times 92 \text{ g}$$

$$m_{\text{etanol, g}} = m_{\text{amido teórico, g}} \times 0,58 \text{ g}$$

Considerando a amostra BR501, cujo teor de amido total foi de 63,38%, essa amostra apresentou massa de 0,6338 g. Logo, para essa cultivar, metanol = 0,360 g em base úmida. A conversão para base seca é feita considerando-se a umidade da amostra. Para a BR501, 10,5% de umidade (89,5% de matéria úmida). Para descontar a umidade, faz-se o cálculo:

$$> \text{massa da amostra} - \text{massa de umidade na amostra} = 0,360 \text{ g} - (0,360 \times 0,105) = 0,320$$

Assim, tem-se que a massa teórica de etanol para a cultivar BR501, é de 0,320 g.

Para fazer a conversão a volume, considerar a densidade do etanol ( $0,7893 \text{ g mL}^{-1}$ ).

- Passo 2: calcular a eficiência

Deve-se ter em mente que a eficiência é a razão entre a quantidade obtida e a esperada. Dessa forma, considerando a amostra BR501 temos que:

$$\text{Eficiência} = \frac{\text{massa obtida de etanol}}{\text{massa teórica, esperada}} = \frac{0,317}{0,320} \times 100 = 99\%$$

<sup>3</sup> Determinado pelo teor de amido total para cada amostra