

UNIVERSIDADE FEDERAL DO AMAZONAS  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

BERNA ALMEIDA DE SOUZA

**AUTENTICAÇÃO DO ÓLEO DE ANDIROBA POR RMN E  
FERRAMENTAS QUIMIOMÉTRICAS**

MANAUS

2019

BERNA ALMEIDA DE SOUZA

**AUTENTICAÇÃO DO ÓLEO DE ANDIROBA POR RMN E  
FERRAMENTAS QUIMIOMÉTRICAS**

Dissertação submetida como parte dos requisitos para obtenção do título de MESTRE EM QUÍMICA pela Universidade Federal do Amazonas, cuja linha de pesquisa é de Química de Produtos Naturais.

Orientador: Sergio Massayoshi Nunomura (COTEI/ INPA)

Coorientador: Luiz Henrique Keng Queiroz Júnior (UFG)

MANAUS

2019

### Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S729a Souza, Berna Almeida De  
Autenticação do óleo de andiroba por RMN e ferramentas  
quimiométricas / Berna Almeida De Souza. 2019  
94 f.: il. color; 31 cm.

Orientador: Sergio Massayoshi Nunomura  
Coorientador: Luiz Henrique Keng Queiroz Júnior  
Dissertação (Mestrado em Química de Prod. Naturais) -  
Universidade Federal do Amazonas.

1. Carapa guianensis. 2. Adulteração. 3. RMN de 1H. 4.  
Quimiometria. I. Nunomura, Sergio Massayoshi. II. Universidade  
Federal do Amazonas III. Título

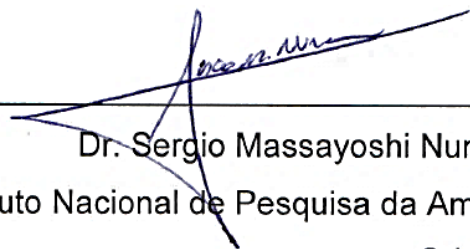
BERNA ALMEIDA DE SOUZA

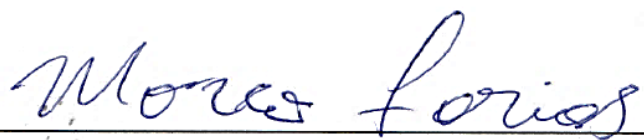
**“AUTENTICAÇÃO DO ÓLEO DE ANDIROBA POR RMN E  
FERRAMENTAS QUIMIOMÉTRICAS”**


Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Química, do Instituto de Ciências Exatas da Universidade Federal do Amazonas como requisito para obtenção do Grau de Mestre em Química.

Aprovada, em 22 de março de 2019.

COMISSÃO EXAMINADORA

  
\_\_\_\_\_  
Dr. Sergio Massayoshi Nunomura  
Instituto Nacional de Pesquisa da Amazônia  
Orientador

  
\_\_\_\_\_  
Dr. Marco Antonio dos Santos Farias  
Membro UFAM

  
\_\_\_\_\_  
Dr.ª Lyege Magalhães Oliveira  
Membro Externo

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus pelo dom da vida.

Aos meus familiares, principalmente a minha mãe, por sempre acreditar e apoiar as minhas escolhas.

Ao Dr. Sergio Massayoshi Nunomura, quem eu admiro muito, pela orientação, pelos conhecimentos compartilhados e pela amizade que contribuíram muito para o meu crescimento profissional e pessoal.

Ao Dr. Luiz Henrique Keng Queiroz Júnior, que gentilmente aceitou o convite de coorientação desse trabalho. Sempre serei grata pela paciência e dedicação em explicar quimiometria e a linguagem do programa R.

Aos técnicos Magno, Sabrina e Zelina do laboratório LTQPN-INPA pelo treinamento do uso do equipamento de RMN e por todos os feriados e finais de semana cedidos para a conclusão dos experimentos de RMN.

À equipe LAPAAM-INPA por toda experiência adquirida e compartilhada durante toda a minha jornada científica.

Aos meus amigos que compartilharam altas doses de cafeína nos momentos de alegria e desespero. Principalmente, a Ritinha que participava de altas discussões sobre quimiometria comigo e geralmente terminávamos com mais dúvidas do que certezas.

Ao programa de pós-graduação pela missão de estudos e por todos os pedidos de prorrogação concedidos.

À família do professor Keng que me acolheu em Goiânia durante a missão de estudos, principalmente à Darlene que sempre foi muito solícita.

Ao INPA onde foram realizados todos os experimentos que foram prescindíveis à conclusão deste trabalho.

À CAPES pelo apoio financeiro.

## RESUMO

O óleo de andiroba é bastante conhecido e comercializado no Brasil e no exterior. A espécie predominante é classificada como *Carapa guianensis*, cuja árvore é considerada nativa da região Amazônica, podendo ser encontrada principalmente na região Norte do país. Grande parte dos estudos de óleo de *C. guianensis*, procurou associar a constituição química da espécie com suas propriedades medicinais, uma vez que a mesma é conhecida por suas propriedades anti-inflamatória, analgésica, inseticida, repelente, antibacteriano, antiparasítico, anticâncer, dentre outras. O óleo de andiroba é constituído em mais de 95% de material graxo (ácido palmítico, oleico, esteárico, linoleico e araquídico) e menos de 5% de material insaponificável, principalmente os limonoides (triterpenos modificados). Por se tratar de um óleo de grande valor medicinal e de não haver controle de qualidade na venda e revenda desse produto, este trabalho visou a utilização da técnica de espectroscopia de RMN de  $^1\text{H}$  1D para desenvolver um método de avaliação e de controle de possíveis adulterações do óleo de andiroba por óleo de soja. A técnica permitiu caracterizar e identificar possíveis adulterações do óleo de andiroba com óleo de soja, em que os dados espectroscópicos foram submetidos a um estudo quimiométrico abrangendo os métodos exploratórios e classificatórios de estatística multivariada. A caracterização da cadeia graxa se deu por meio dos experimentos de RMN de  $^1\text{H}$  1D e 2D (COSY), bem como experimentos heteronucleares 2D (HMBC e HSQC). Tais experimentos, confirmaram as características da cadeia graxa de ambos os óleos vegetais. O óleo de andiroba possui um percentual muito elevado de compostos saturados e de ácido oleico, quando existe a presença de soja em amostras de andiroba, a adulteração é observada porque aumenta a área dos sinais referentes aos insaturados, além disso a soja também possui o ácido linolênico que não está presente no óleo de andiroba. Essas informações foram úteis para interpretar o comportamento dos dados espectroscópicos por meio da quimiometria. Os métodos exploratórios aplicados nesse estudo avaliaram o comportamento e as tendências das amostras pela Análise de Agrupamento Hierárquico (HCA), *Fuzzy C-Means* (FCM) e Análise de Componentes Principais (PCA). A partir dos métodos exploratórios foi possível realizar o método classificatório de regressão dos mínimos quadrados parciais por intervalos (I-PLS), em que foi proposto um modelo para estimar o quanto de adulteração estava presente em cada amostra comercial adquirida nos comércios e feiras da cidade de Manaus. Para isso, construiu-se um banco de dados de adulteração conhecida, cuja combinação binária do óleo padrão de andiroba mais óleo de soja foram utilizados para validar o método de calibração. Além de classificar as amostras em íntegras e adulteradas, também foi possível estimar a adulteração detectada nas amostras comerciais.

Palavras-chave: *Carapa guianensis*. Adulteração. RMN de  $^1\text{H}$ . Quimiometria.

## ABSTRACT

Andiroba oil is widely known and traded in Brazil and abroad. The most abundant andiroba species is *Carapa guianensis*, which is considered native from the Amazon region, and can be found mainly in the northern region of the country. Most of the studies of *C. guianensis* oil are related to its chemical constitution, along with their medicinal properties, especially anti-inflammatory, analgesic, insecticidal, repellent, antibacterial, antiparasitic and anticancer properties. Up to 95% of andiroba oil is constituted by fatty material (palmitic, oleic, stearic and arachidic acid) and less than 5% of unsaponifiable matter, mostly limonoids. Due to its great commercial and medicinal value, this work aimed to employ one-dimensional  $^1\text{H}$  NMR spectroscopy for the evaluation and control of adulteration of andiroba by soybean oil. This technique allowed to characterize and to identify adulterations of andiroba oil with soybean oil identified with the help of chemometric study employing both exploratory and classificatory methods. The characterization of the fatty chain of the andiroba and soybean oil was determined by 1D and 2D hydrogen NMR experiments (COSY), as well as 2D heteronuclear experiments (HMBC and HSQC). These experiments confirmed the different characteristics of the fatty chain of both vegetable oils. Andiroba oil has a very higher percentage of saturated chains and oleic acid, whereas soybean oil a higher proportion of unsaturated chains. The increasing area of the unsaturated hydrogens peaks in the  $^1\text{H}$  NMR spectra help to identify the adulteration of andiroba by soybean oil. In addition, the soybean also contains the linolenic acid that is not present in the andiroba oil. The  $^1\text{H}$  NMR data were analyzed by statistical exploratory methods to evaluate the behavior and trends of the samples by Hierarchical Cluster Analysis (HCA), *Fuzzy C-Means* and Principal Component Analysis (PCA). From the exploratory results it was possible to perform the interval least squares regression (I-PLS) method, in which a model was proposed to estimate the adulteration in each commercial sample acquired in the market and fairs of the city of Manaus. Firstly, a known adulteration database was built to validate the calibration method, from a combination of standard andiroba oil with increasing amounts of soybean oil. This enabled us to classify the samples as non-adulterated and adulterated, and also to estimate the adulteration detected in commercial samples.

## LISTA DE FIGURAS

Figura 1: Distribuição geográfica das espécies de <i>C. guianensis</i> (A) e <i>C. procera</i> (B) .....	19
Figura 2: Sementes de andiroba das espécies <i>C. guianensis</i> (A) e <i>C. procera</i> (B)...	20
Figura 3: Principais componentes de formação dos triacilglicerídeos (TAG) .....	24
Figura 4: Ácidos graxos livres que fazem parte da composição da cadeia graxa do óleo de andiroba.....	25
Figura 5: Alguns limonoides isolados de <i>C. guianensis</i> .....	26
Figura 6: Região dos hidrogênios beta da unidade do glicerol antes e após o alinhamento no programa MATLAB através do algoritmo ICOSHIFT. ....	41
Figura 7: Espectro de RMN de <sup>1</sup> H dos óleos vegetais e as regiões de exclusão representadas pela cor rosa.....	41
Figura 8: Representação dos <i>buckets</i> automáticos cortando as áreas de sinais do espectro de RMN de <sup>1</sup> H dos óleos vegetais .....	42
Figura 9: Representação dos <i>buckets</i> após a otimização manual .....	43
Figura 10: Espectros de RMN <sup>1</sup> H destacando os tipos de hidrogênios presentes nos óleos de soja, andiroba padrão e adulterada, em que R são os ésteres característicos das amostras.....	45
Figura 11: Mapa de contornos COSY adquirido para a amostra de soja (S02) .....	47
Figura 12: Ampliação do mapa de contornos COSY da região do grupo metila do éster linolenato .....	48
Figura 13: Mapa de contornos do experimento HSQC do óleo de soja (S02).....	49
Figura 14: Mapa de contornos do experimento HMBC do óleo de soja (S02) .....	50
Figura 15: Espectro médio de RMN de <sup>1</sup> H das amostras do banco 1 (andiroba padrão x andiroba comercial x soja) com a representação dos <i>buckets</i> otimizados, matriz 63x31 .....	53
Figura 16: Dendrograma da HCA das amostras de andiroba padrão e andiroba comercial.....	55
Figura 17: Dendrograma da HCA das amostras andiroba padrão e andiroba comercial <i>versus</i> soja comercial .....	57
Figura 18: Mapa de calor da média das réplicas das amostras de andiroba padrão e comercial <i>versus</i> o óleo de soja comercial.....	58



Figura 19: Representação gráfica em silhueta da avaliação do melhor agrupamento encontrado $K = 3$ .....	59
Figura 20: Representação gráfica da análise de <i>Fuzzy C-Means</i> da média das réplicas das amostras do banco 1 .....	59
Figura 21: Gráficos de scores e Loadings PC1 X PC2 da PCA (SVD).....	61
Figura 22: Gráfico de desempenho das variâncias das componentes principais (PC1 a PC5) e suas contribuições para a PCA robusta .....	62
Figura 23: Gráfico de <i>scores</i> e <i>loadings</i> PC1 (37,63%) versus PC2 (30,64%) da PCA (Robusta) das amostras de óleo de andiroba.....	64
Figura 24: Sinais de RMN de $^1\text{H}$ das amostras de óleo de andiroba padrão (verde), andiroba comercial íntegra (laranja), andiroba comercial adulterada (vermelho) e soja (roxo). Os números acima dos sinais são os números dos buckets encontrados nos gráficos de Loadings .....	66
Figura 25: Gráficos de <i>scores</i> e <i>loadings</i> PC2 (30,64%) versus PC3 (21,81%) das amostras de óleo de andiroba padrão, andiroba comercial adulterada, andiroba comercial íntegra e soja .....	68
Figura 26: Gráfico de avaliação do erro médio quadrático de predição versus variáveis latentes (componentes) .....	70
Figura 27: Correlação entre os valores de soja predito ( <i>predict</i> ) pelo modelo e adicionados às misturas ( <i>measured</i> ).....	71
Figura 28: Gráfico de seleção de modelo.....	71
Figura 29: Correlação entre os valores de soja predito ( <i>predict</i> ) pelo modelo e adicionados às misturas ( <i>measured</i> ) após a validação do modelo com 2 componentes. ....	73

## LISTA DE QUADROS

Quadro 1: Informações dos valores comerciais atribuídos ao óleo e semente de andiroba de algumas lojas físicas e virtuais no Brasil no ano de 2018 .....	23
Quadro 2: Técnicas aplicadas a análise de óleos e gorduras vegetais .....	28
Quadro 3: Amostras de óleo de andiroba utilizadas nos experimentos de RMN.....	35
Quadro 4: Amostras do <b>Banco de dados 1</b> para análise quimiométrica não supervisionada .....	37
Quadro 5: Matriz de investigação ou matriz Y apresenta as amostras de concentração desconhecida de possível adulteração com óleo de soja .....	39

**LISTA DE TABELAS**

Tabela 1: Matriz de supervisão ou matriz X apresenta o grau adulteração do óleo de andiroba com óleo de soja .....	38
Tabela 2: Intervalos de frequência das regiões de exclusão dos espectros de RMN de <sup>1</sup> H dos óleos vegetais.....	41
Tabela 3: Composição da cadeia graxa dos óleos de andiroba e soja por cromatografia gasosa (CG).....	44
Tabela 4: Atribuição dos deslocamentos químicos dos hidrogênios de triacilgliceróis dos óleos vegetais por RMN de <sup>1</sup> H unidimensional.....	46
Tabela 5: Atribuição dos deslocamentos químicos dos carbonos de triacilgliceróis dos óleos vegetais por meio dos experimentos de HSQC e HMBC.....	51
Tabela 6: Misturas ( <i>blends</i> ) utilizadas na construção do modelo de calibração e validação .....	72
Tabela 7: Resultados de concentração (% , v/v) de óleo de soja em amostras comerciais pelo método I-PLS.....	74

## LISTA DE ABREVIATURAS E SIGLAS

- 1D – Unidimensional
- 2D – Bidimensional
- AG – Ácidos graxos
- AGI – Ácidos graxos insaturados
- AGS – Ácidos graxos saturados
- AQ – Tempo de aquisição entre cada varredura (*Acquisition time*)
- CG – Cromatografia gasosa
- CG-DIC – Cromatografia Gasosa acoplada ao Detector de Ionização de Chama
- CG-EM – Cromatografia Gasosa acoplada Espectrometria de Massas
- CLAE – Cromatografia Líquida de Alta Eficiência
- CLAE-DAD – Cromatografia Líquida de Alta Eficiência acoplada ao Detector Arranjo de Diodos
- CLAE-UV – Cromatografia Líquida de Alta Eficiência acoplada ao Ultravioleta
- CDCl<sub>3</sub> – Clorofórmio deuterado
- COSY – *CORrelation SpectroscopY*
- DAD – Detector Arranjos de Diodos
- DIC – Detector de Ionização de Chamas
- DOSY – *Diffusion Ordered SpectroscopY*
- EM – Espectrometria de Massas
- EMA – Adulteração Economicamente Motivada (*Economically motivated adulteration*)
- FTIR-HATR – Espectroscopia por Refletância Total Atenuada Horizontal no Infravermelho com Transformada de Fourier
- FT-NIR – Transformada de Fourier-Espectroscopia no Infravermelho Próximo
- FCM - *Fuzzy C-Means* (Análise de Cluster Nebulosa)
- HMBC – *Heteronuclear Multiple Bond Correlation*
- HSQC – *Heteronuclear Single Quantum Correlation*
- lb – *Line Broadening*
- INPA – Instituto de Pesquisas da Amazônia
- J-RES – *J-Resolved Spectroscopy*
- LAPAAM – Laboratório de Princípios Ativos da Amazônia
- NS – Número de promediações (*number of scans*)
- P1 – Pulso (*Pulse*)

PCA – Análise de Componentes Principais (*Principal Component Analysis*)

PCAPP – Robust PCA by Projection Pursuit.

I-PLS – Regressão por Mínimos Quadrados Parciais por Intervalos (*Interval Partial Least Square*)

ppm – Partes Por Milhão

RMN – Ressonância Magnética Nuclear

RMSEP – Raiz Quadrada do Erro Médio Quadrático da Predição (*Root Mean Square Error of Prediction*)

SI – Número de pontos reais no espectro (*Size of real spectrum*)

SIMCA – Soft Independent Modeling of Class Analogy

SWH – Largura da janela espectral (*Spectral Width*)

T1 – Tempo de relaxação longitudinal

TAG – TriAcilGlicerídeo

TMS – TetraMetilSilano

UV-Vis – UltraVioleta Visível

v/v – Volume por Volume

## LISTA DE SÍMBOLOS

$^1\text{H}$  – Prótio

$^2\text{H}$  – Deutério

$^7\text{Li}$  – Lítio

$^{11}\text{B}$  – Boro

$^{13}\text{C}$  – Carbono

$^{14}\text{N}$  – Nitrogênio

$^{19}\text{F}$  – Flúor

$^{31}\text{P}$  – Fósforo

% – Porcentagem

$\delta$  – Deslocamento químico

$\mu\text{L}$  – Microlitro ( $\mu = 10^{-3}$ )

d – Duplete

Hz – Hertz

$J$  – Constante de Acoplamento

K – Kelvin

MHz – Megahertz

R\$ – Moeda brasileira (real)

s – Singleto

t – Tripleto

## SUMÁRIO

1. INTRODUÇÃO .....	16
2. DESENVOLVIMENTO .....	19
2.1 Andiroba.....	19
2.1.1 Utilização e comercialização da andiroba.....	21
2.1.2 Composição química do óleo de Andiroba .....	24
2.2 Adulteração de óleos vegetais .....	26
2.2.1 Técnicas aplicadas à adulteração de óleos vegetais.....	27
2.2.2 RMN aplicada à adulteração de óleos .....	30
2.3 Quimiometria e algumas fundamentações .....	31
3. OBJETIVOS.....	34
3.1 Objetivo geral .....	34
3.2 Objetivos específicos .....	34
4. METODOLOGIA .....	35
4.1 Obtenção dos óleos de andiroba .....	35
4.2 Materiais e equipamentos .....	36
4.3 Obtenção dos espectros bidimensionais para caracterização dos óleos .....	36
4.4 Preparo das amostras do Banco de dados 1 (andiroba padrão, andiroba comercial e soja) para análise quimiométrica não supervisionada .....	37
4.5 Preparo das amostras do Banco de dados 2 para análise quimiométrica supervisionada.....	38
4.6 Aquisição dos espectros de RMN de $^1\text{H}$ para análise quimiométrica.....	39
4.7 Análise quimiométrica dos Banco de dados 1 e Banco de dados 2.....	40
4.7.1 Pré-tratamento dos dados no programa TopSpin de RMN de $^1\text{H}$ para análise quimiométrica .....	40
4.7.2 Montagem da Matriz, pré-tratamento e pré-processamento dos dados no MATLAB e RStudio.....	40
5. RESULTADOS E DISCUSSÃO .....	44

5.1	Caracterização química dos óleos vegetais .....	44
5.1.1	Atribuição dos sinais nos espectros de RMN de $^1\text{H}$ unidimensional dos óleos de andiroba e soja.....	44
5.1.2	Atribuição dos sinais nos espectros de RMN de $^1\text{H}$ bidimensional homonuclear (COSY) dos óleos de soja e andiroba.....	47
5.1.3	Atribuição dos sinais nos espectros de RMN bidimensional heteronuclear (HSQC e HMBC) dos óleos de soja e andiroba.....	49
5.2	Análise multivariada dos dados.....	51
5.2.1	Análise exploratória ou não supervisionada .....	54
5.2.1.1	Análise de agrupamento hierárquico (HCA) do Banco de dados 1 (Andiroba padrão, andiroba comercial e soja) .....	54
5.2.1.2	Análise de agrupamento <i>Fuzzy C-Means</i> (FCM) do Banco de dados 1 (andiroba padrão, andiroba comercial e soja) .....	58
5.2.1.3	Análise exploratória por PCA robusta do Banco de dados 1 (andiroba padrão, andiroba comercial e soja) .....	60
5.2.2	Análise classificatória ou supervisionada pelo método dos mínimos quadrados parciais (I-PLS) do Banco de dados 2. ....	69
5.2.2.1	Modelo de calibração.....	69
5.2.2.2	Validação do Modelo de I-PLS .....	72
5.2.2.3	Teste com as amostras comerciais de andiroba.....	73
6.	CONCLUSÃO .....	75
7.	REFERÊNCIAS .....	77
8.	ANEXOS.....	83



## 1. INTRODUÇÃO

O Brasil é conhecido mundialmente por suas riquezas naturais, principalmente quando se trata da flora Amazônica. O uso etnofarmacológico de plantas medicinais é bastante difundido em comunidades e grupos étnicos da região Norte do país (VIEGAS JR. e BOLZANI, 2006). Dentre a vasta flora que predomina na região, o óleo de andiroba apresenta grande potencial terapêutico, como: antisséptico, anti-inflamatório, inseticida, antipirético e repelente de mosquitos do gênero *Anopheles* e da espécie *Aedes aegypti*, que são responsáveis pela transmissão das doenças tropicais, como: malária, dengue, chicungunha e zika (MIRANDA JR *et al.*, 2012).

A andiroba pertence à família Meliaceae, cuja espécie predominante no Amazonas é a *Carapa guianensis*, podendo também ser encontrada a espécie *C. procera*, ambas são conhecidas popularmente por andiroba e de difícil diferenciação. A extração do óleo das sementes é de grande importância para comunidades ou cooperativas da zona rural da região Norte do país.

Atualmente não existe um controle da origem, qualidade e adulteração da composição química do óleo de andiroba comercial, facilitando a sua adulteração com a mistura com óleos mais baratos, como é o óleo de soja, com o objetivo de aumentar o lucro. Como a composição da cadeia graxa do óleo de andiroba é semelhante ao óleo de soja, a adulteração tem se intensificado bastante, tanto nas regiões de exploração como na revenda deste produto a terceiros. Essa falta de controle de qualidade tem desvalorizado o mercado do óleo de andiroba, apesar de ser bastante valorado pelas suas propriedades medicinais (CABRAL *et al.*, 2013; GASPAR AMARAL e FIERRO, 2013).

Existem vários trabalhos com óleos vegetais que avaliam o perfil químico afim de encontrar informações que possam identificar a natureza da amostra quanto a sua autenticação e possíveis adulterações. O perfil químico apresenta a composição das substâncias presentes na amostra que são denominadas como impressão digital (*fingerprint*). Para se obter um perfil químico são utilizadas diversas técnicas instrumentais que podem ser realizadas individualmente ou em conjunto, a fim de tentar obter a maior quantidade de informação possível da composição química. Dentre as técnicas que são abordadas para obtenção de perfil químico podem ser encontradas as cromatográficas, espectrométricas e espectroscópicas, das quais avaliam desde a estrutura das substâncias quanto suas concentrações. Apesar dos

avanços instrumentais é impossível caracterizar cada substância presente em uma só análise e em tão pouco tempo utilizando apenas uma técnica analítica. Contudo, é possível conseguir informações de caráter singular sobre determinada amostra dependendo do objetivo do estudo, como por exemplo, direcionar o estudo a uma classe de substâncias características dessa amostra (*targed analyses*) (LAMANNA, 2013).

Em óleos vegetais, também conhecidos por óleos fixos, a classe predominante na sua composição são os lipídeos, cujas substâncias majoritárias são os triacilglicerídeos que são derivados de ácidos graxos. Para esse tipo de substância a técnica de análise mais utilizada é a cromatografia gasosa acoplada ao detector de ionização de chama (GC-FID) ou espectrômetro de massas (GC-MS). Apesar da técnica ser bastante consolidada para avaliação da cadeia graxa de óleos vegetais fixos, existe uma crescente utilização da técnica de espectroscopia de ressonância magnética nuclear (RMN), especialmente os experimentos de RMN de  $^1\text{H}$  unidimensional, cujo experimento necessita de preparo mínimo da amostra, aquisição dos dados em um curto espaço de tempo e, dependendo da aquisição pode ser utilizada para fins qualitativos e quantitativos. Por muito tempo, a técnica de RMN foi bastante utilizada para elucidar estruturas químicas em estudos com produtos naturais, porém existem muitos estudos empregando RMN no controle de qualidade, na avaliação de reações, quantificação de metabólitos contaminantes ou marcadores químicos de autenticação, dentre outros (BARISON *et al.*, 2010; RUIZ-SAMBLÁS, *et al.* 2012; CABRAL, 2013).

Assim como toda técnica instrumental, os experimentos de RMN de  $^1\text{H}$ , principalmente voltados ao controle de qualidade, também geram uma quantidade de resposta volumosa, por envolver amostragens específicas como réplicas e coletas<sup>1</sup> diferentes. Nesses casos, torna-se necessário o uso da quimiometria que auxilia matematicamente e estatisticamente a avaliação do comportamento desses conjuntos de dados de forma exploratória ou classificatória. Como são realizadas várias medições, as análises estatísticas multivariadas são as mais utilizadas em quimiometria, podendo ser denominadas não supervisionada ou supervisionada. Normalmente, iniciam-se os estudos com as análises não supervisionadas ou exploratórias, pois elas indicam como são os comportamentos das amostras sem

---

<sup>1</sup> Coletas também pode ser referente a compra do mesmo produto só que de lotes, marcas ou locais diferentes.

nenhum conhecimento prévio. Dentre as mais utilizadas, encontram-se a Análise de Componentes Principais (PCA) e Análise de Agrupamento Hierárquico (HCA). Em estudos supervisionados ou classificatórios os métodos mais utilizados são os de calibração e regressão pelo método dos mínimos quadrados (I-PLS) e método do vizinho mais próximo (KNN), estes avaliam a similaridade das amostras com uma matriz de supervisão.

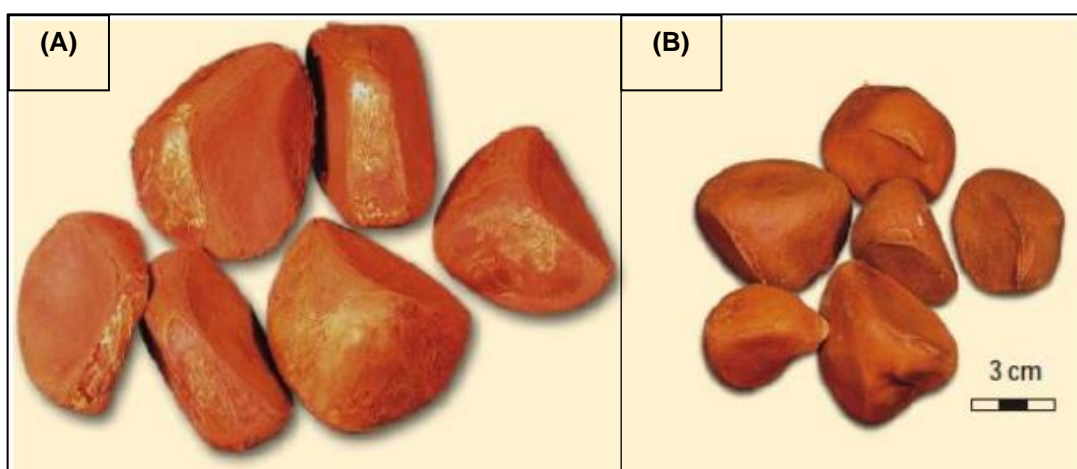
Devido ao grande crescimento da utilização dos experimentos de RMN de  $^1\text{H}$  para controle de qualidade e identificação de adulteração em óleos vegetais, realizou-se um estudo quimiométrico de dados obtidos por RMN de  $^1\text{H}$  unidimensional de amostras de óleo de andiroba adulterados com óleo de soja. Essa técnica é capaz de realizar estudos qualitativos e quantitativos que podem ser facilmente reproduzidos, o que é pertinente para o desenvolvimento de análises empregadas no controle de qualidade e que podem ser utilizadas para valorizar o produto perante o mercado consumidor.



A *C. guianensis* é uma árvore de médio a grande porte, chegando a atingir de 10 a 30 m de altura, com diâmetro acima do peito de 15 a 30 cm; as folhas medem de 30 a 70 cm de comprimento; o fruto é globuloso, de cor castanha e possui 7-9 x 7-10 cm. Esse fruto é dividido por valvas e cada uma pode conter até 5 sementes de 2,4-4,5 x 2,3-5,1 cm. A andiroba floresce uma vez por ano, entre os meses de agosto e outubro, e frutifica de janeiro a maio podendo variar de acordo com a região, bem como as dimensões desse indivíduo. A *C. procera* ocorre como uma árvore que pode atingir 30 m de altura. Seus frutos são da forma de ouriço com 4 a 6 valvas, e cada valva pode conter de 1 a 4 sementes. O tamanho de suas sementes é bem menor quando comparada às sementes de *C. guianensis*. (FERRAZ *et al.*, 2002; CAVALCANTE *et al.*, 1986; PENNINGTON *et al.*, 1981).

Na região Amazônica, as duas espécies se desenvolvem em várzeas, terra firme e igapó. Geralmente, as árvores de terra firme são consideradas como uma árvore de grande porte e que produz uma excelente madeira de cor avermelhada para fins comerciais. Já no Norte da Guiana, as andirobeiras crescem em áreas que são inundadas pelas marés, cujas árvores são de médio porte, a cerca de 10 m de altura e apresentam uma madeira esbranquiçada (FISCH *et al.*, 1995; FERRAZ, 2002; KENFACK, 2011a). As espécies *C. guianensis* e a *C. procera* são difíceis de diferenciar botanicamente, apesar disso as pequenas variações morfológicas dos frutos, sementes, flores e folhas podem auxiliar na sua identificação.

Figura 2: Sementes de andiroba das espécies *C. guianensis* (A) e *C. procera* (B)



Fonte: FERRAZ, 2002.

Estudos de morfologia das duas espécies predominantes na região Amazônica se basearam fortemente no número de sementes por lóculo para justificar que o número era contínuo. Contudo, o número de sementes não é confiável devido ao aborto de alguns óvulos ou à compressão de algumas sementes durante o desenvolvimento do fruto. No entanto, a cor e adornos dos frutos e sementes, assim como o comprimento do hilo em relação ao comprimento da semente são importantes na discriminação morfológicas das espécies (KENFACK, 2011a).

### 2.1.1 Utilização e comercialização da andiroba

A andiroba fornece madeira de excelente qualidade que é utilizada na construção civil, fabricação de móveis e acabamentos internos de embarcações. Atualmente, não é realizada com muita frequência a exploração madeireira da andiroba, pois programas de reflorestamento apoiam a preservação e conservação da Floresta Amazônica, visando a comercialização de produtos não madeireiros. Além disso, o principal interesse pela andiroba se dá pelo óleo obtido de suas sementes (FERRAZ *et al.*, 2002; SHANLEY, 2005; DA SILVA *et al.*, 2010).

O manejo sustentável das andirobeiras, coleta das sementes e a produção do óleo andiroba é realizado pelas comunidades extrativistas da região Norte do país, abrangendo os estados do Amazonas e Pará. Geralmente, o óleo obtido nas comunidades extrativistas é utilizado pela própria população local ou vendido em feiras das cidades vizinhas. As comunidades extrativistas costumam utilizar o óleo como relaxante muscular, cicatrizante de feridas, anti-inflamatório contra garganta inflamada e repelente de insetos. Através do óleo também fabricam sabonetes e velas inseticidas (FERRAZ, 2002; DA SILVA *et al.*, 2010; CABRAL, 2013; LIRA-GUEDES e NARDI, 2015). Suas propriedades organolépticas são bem características, pois apesar do óleo apresentar sabor extremamente amargo, o odor que ele exala é adocicado; tem uma coloração amarelada e quando sujeito a temperatura igual ou abaixo de 25 °C, ele se solidifica.

As indústrias farmacêuticas utilizam o óleo de andiroba puro ou associado a outro produto obtido da Floresta Amazônica. Geralmente, fabricam óleos ou cremes de massagem e xaropes. Já as indústrias cosméticas, além de óleos e cremes de massagem, também fabricam cremes hidratantes, xampu, condicionador, perfumes e sabonetes. No Brasil, os produtos industrializados são produzidos e comercializados

pelas empresas Natura, Apuá, Amazon Ervas, Pharmakos da Amazônia e Fitobel (GUEDES, 2008; DA SILVA *et al*, 2010).

Conforme publicado no Diário Oficial da União (2018), o MAPA (Ministério de Estado da Agricultura, Pecuária e Abastecimento) ajustou o preço mínimo dos produtos extrativos da safra de 2018, onde o valor mínimo atribuído à amêndoa de andiroba é de R\$ 1,60/ Kg. Contudo, por ser um produto de origem amazônica, o valor agregado ao óleo e a semente de andiroba é bastante superior quando comercializado por vendedores secundários, conforme descrito no quadro 1. Os valores destacados no quadro 1 foram levantados em algumas lojas virtuais do Brasil e físicas da cidade de Manaus-AM.

Quadro 1: Informações dos valores comerciais atribuídos ao óleo e semente de andiroba de algumas lojas físicas e virtuais no Brasil no ano de 2018

Empresa	Produto	Valor, R\$	Quantidade	Fonte
Natura	Óleo concentrado	59,50	30 mL	<a href="https://sites.natura.com.br/revista/espaco-natura-c14-v1">https://sites.natura.com.br/revista/espaco-natura-c14-v1</a>
	Óleo de massagem	51,60	100 mL	
Amazon Produtos	Óleo puro	104,90	1000 mL	<a href="http://www.amazonprodutos.com.br/nossos-produtos/oleos-naturais.html">http://www.amazonprodutos.com.br/nossos-produtos/oleos-naturais.html</a>
Amazon Ervas	Óleo de massagem	10,00	30 mL	Loja física, Manaus
Pharmakos D' Amazônia	Óleo de massagem	12,00	30 mL	Loja física, Manaus
Pronatus	Óleo de massagem	10,00	30 mL	<a href="http://pronatus.com.br/loja/inicio/38-oleo-de-amendoas.html">http://pronatus.com.br/loja/inicio/38-oleo-de-amendoas.html</a>
Bela Vida Natural	Óleo extra virgem	145,00	1 Kg	<a href="https://www.belavidanatural.com.br/p/oleo-de-andiroba-amazonia-1kg">https://www.belavidanatural.com.br/p/oleo-de-andiroba-amazonia-1kg</a>
Terra Flor	Óleo prensado a frio	55,00	60 mL	<a href="https://terra-flor.com/loja/oleos-vegetais/andiroba-60ml/">https://terra-flor.com/loja/oleos-vegetais/andiroba-60ml/</a>
Phytoterápica	Óleo extra virgem	38,00	60 mL	<a href="https://www.phytoterapica.com.br/loja/oleos-vegetais/oleo-vegetal-de-andiroba-60ml">https://www.phytoterapica.com.br/loja/oleos-vegetais/oleo-vegetal-de-andiroba-60ml</a>
Ecycle		45,00	100 mL	<a href="https://www.ecycle.eco.br/oleo-andiroba-poliioleos.html">https://www.ecycle.eco.br/oleo-andiroba-poliioleos.html</a>
MFRural	Sementes	55,00	1 Kg	<a href="http://www.mfrural.com.br/detalhe/oleo-andiroba-252090.aspx">http://www.mfrural.com.br/detalhe/oleo-andiroba-252090.aspx</a>

\*Dados coletados durante o ano de 2018.



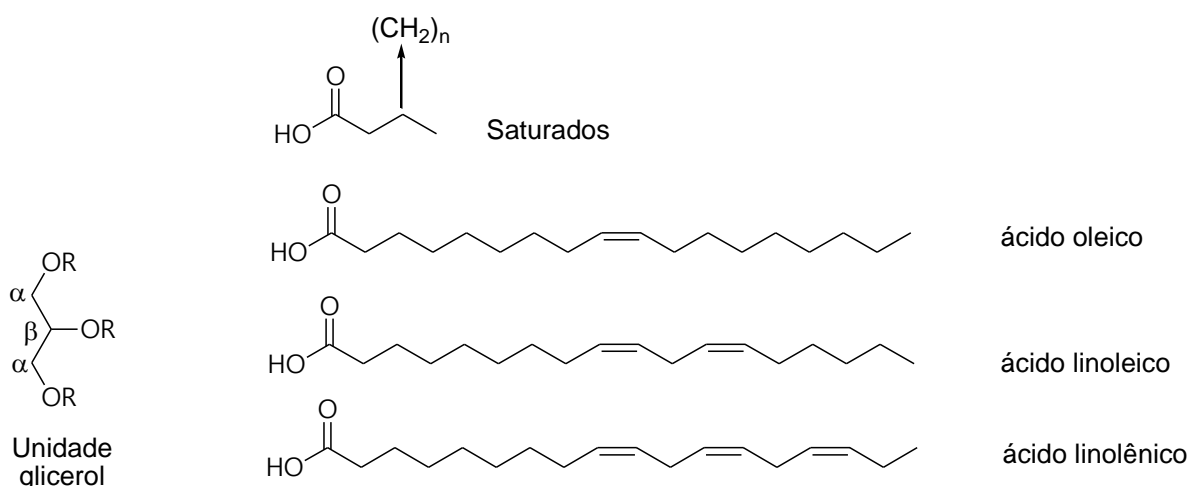
## 2.1.2 Composição química do óleo de Andiroba

O óleo de andiroba, assim como todo óleo vegetal e gordura animal é composto principalmente por triacilglicerídeos, abrangendo cerca de 95-98% da sua composição química. Os triacilglicerídeos (TAGs) são ésteres formados, a partir de três moléculas de ácidos graxos (AG) com uma molécula de glicerol (carboidrato), como representado na figura 3 (INDELICATO *et al*, 2017; REZANKA *et al*, 2017).

A composição dos óleos em geral é distinguida de acordo com a estrutura dos TAGs, no qual compreende o tamanho da cadeia, posição de ligação dupla e configuração *cis* / *trans*; a distinção também é correlacionada à proporção de ácidos graxos saturados (AGS) e insaturados (AGI), destacando o número e posição das ligações duplas na cadeia carbônica (SOMPILA *et al.*, 2017).

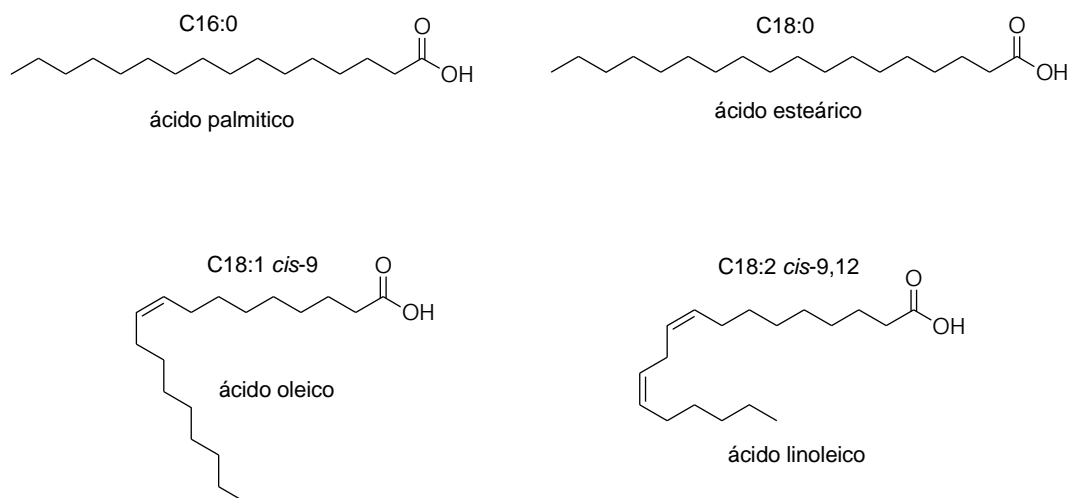
A cadeia graxa do óleo de andiroba é constituída predominantemente por ácido oleico (51,81 %), palmítico (25,76 %), esteárico (9,08 %), linoleico (8,3 %), figura 4 (CABRAL, 2013; DE SANTANA, 2018). Apesar dos ácidos graxos insaturados (AGI) prevalecerem na composição do óleo de andiroba, dependendo do local e temperatura ao qual é armazenado, o óleo pode ter aspecto de gordura (semissólido ou sólido) devido a presença de 34,84% de AGS.

Figura 3: Principais componentes de formação dos triacilglicerídeos (TAG)



Fonte: Adaptado de BARISON, 2010.

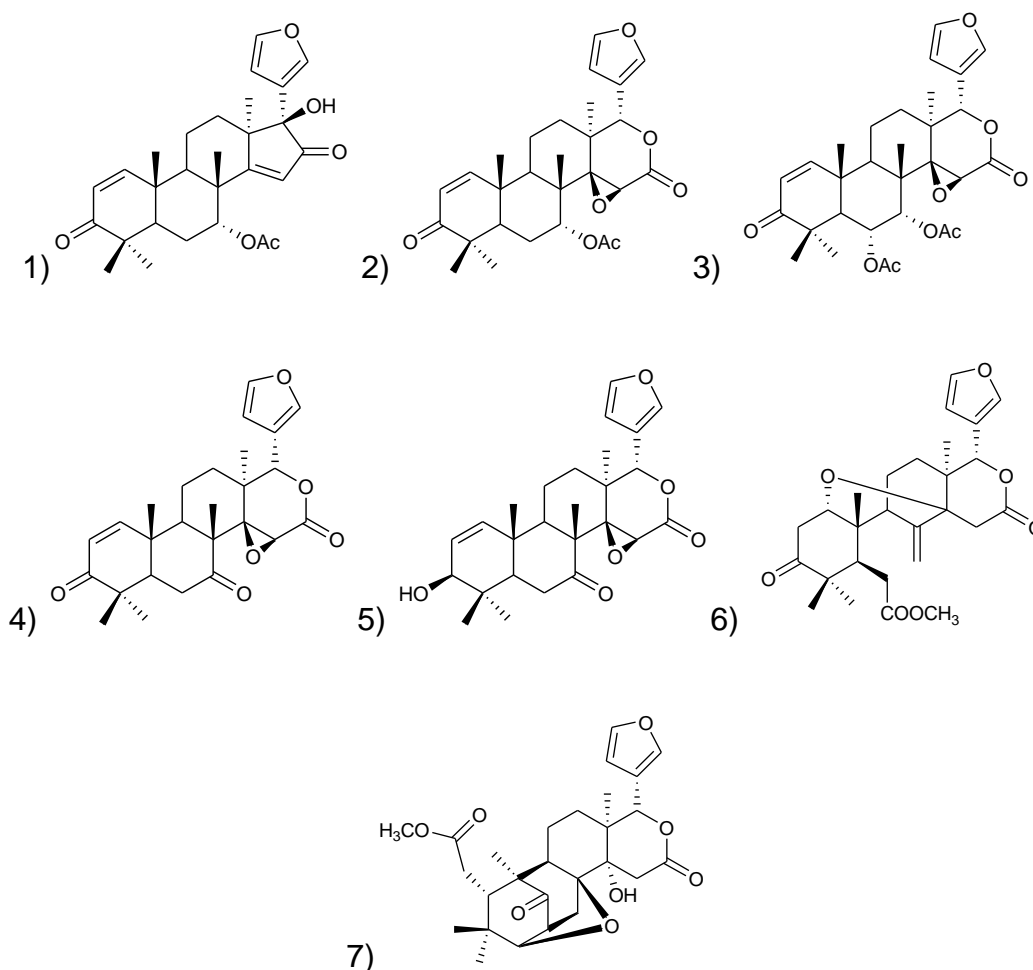
Figura 4: Ácidos graxos livres que fazem parte da composição da cadeia graxa do óleo de andiroba



As propriedades medicinais são comumente associadas aos metabólitos secundários presentes no material insaponificável do óleo, que perfazem cerca de 2 a 5% do óleo e que compreendem triterpenos, esteroides, carotenoides, cumarinas, flavonoides e limonoides. Dentre esses, a maioria dos estudos científicos atribuem a atividade biológica aos limonoides, principalmente a inseticida e anti-inflamatória. Existem relatos que alguns grupos indígenas e a população das comunidades extrativistas já utilizavam o óleo para combater a ação de mosquitos, através da fabricação de velas ou emulsões. Outras propriedades também são associadas ao uso do óleo das sementes de andiroba como analgésico, anti-inflamatório, antibacteriano e antiparasítico (FERRAZ, 2002; AMBROSIN, 2006; CABRAL, 2103).

Os limonoides são triterpenóides com alto grau de oxigenação derivados de um precursor com um esqueleto furanilesteróide que está sujeito a reações de rearranjo. Essas reações podem gerar modificações na estrutura básica ocasionando alterações expressivas nas atividades biológicas dos limonoides. Dentre os limonoides isolados e identificados em *C. guianensis*, encontram-se a 17 $\beta$ -hidroxiazadiradiona (1), gedunina (2), 6 $\alpha$ -acetoxi-gedunina (3), 7-desacetoxi-7-oxogedunina (4), 1,2-dihidro-3 $\beta$ -hidroxi-7-deacetoxi-7-oxogedunina (5), angolensato de metila (6), e xiloccensina k (7), acompanhar na figura 5 (AMBROSIN, 2006; SILVA *et al.*, 2012; CABRAL, 2013; DE SANTANA, 2018).

Figura 5: Alguns limonoides isolados de *C. guianensis*



Fonte: CABRAL, 2013.

## 2.2 Adulteração de óleos vegetais

Os óleos vegetais podem ser utilizados em distintas aplicações, desde a alimentícia quanto a combustíveis. Devido a sua multifuncionalidade, também são altamente procurados por indústrias farmacêuticas, cosméticas e agrônômicas. Várias espécies vegetais são utilizadas no Brasil e no mundo na produção de óleos vegetais com uma forma de produção extensiva e mecanizada como é o caso da soja, palma, canola, do milho, algodão e girassol. Na Amazônia, a produção é em escala muito menor e baseada no manejo, a partir de espécies nativas em adensamentos florestais naturais com destaque para os óleos de andiroba, buriti, copaíba e o pequi (SANTOS, 2018; CONAB, 2019).

As características química e física do óleo vegetal, bem como suas propriedades funcionais são associadas ao período de colheita, condições climáticas,

região de cultivo e método de extração. Tais parâmetros podem ser correlacionados como controle de qualidade dos óleos vegetais. Vários estudos sobre o azeite de oliva tiveram como objetivo determinar a origem, a qualidade e adulteração deste óleo, por meio de técnicas analíticas associadas a análise quimiométrica dos dados obtidos.

Devido ao seu alto valor agregado, o azeite de oliva é frequentemente adulterado com óleo de bagaço de azeitona, com óleo residual do refino alcalino, alguns óleos de composição semelhante (óleo de avelã e de girassol com alto teor de ácido oleico) ou óleos com preço mais em conta (palma, abacate e canola). Essa atividade é definida como Adulteração Economicamente Motivada (*Economically motivated adulteration, EMA*). As adulterações podem ser identificadas como práticas intencionais de diluição ou substituição, contaminação, aprimoramento artificial, falsificação, remoção de constituintes autênticos, alegação fraudulenta no rótulo (TSIMIDOU, *et al.* 2016).

### 2.2.1 Técnicas aplicadas à adulteração de óleos vegetais

Com o desenvolvimento das técnicas analíticas está sendo possível realizar estudos diferenciados que realmente alcancem avaliações qualitativas e/ou quantitativas da maioria dos metabólitos de um organismo. Com isso facilita o mapeamento e identificação de marcadores químicos ou substâncias bioativas em quantidade mínima de material, evitando-se a marcha analítica tradicional da fitoquímica, na qual envolve várias etapas exaustivas de fracionamento e isolamento de substâncias (FUNARI *et al.*, 2013). Essas técnicas instrumentais são bastante empregadas em estudos de adulteração de óleos vegetais, apesar da caracterização de óleos vegetais ter sido bastante consolidada pela cromatografia gasosa acoplada ao detector de ionização de chama (CG-DIC) e espectrometria de massas (CG-EM) (RUIZ-SAMBLÁS, *et al.* 2012; CABRAL, 2013), ver quadro 2. Ambas as técnicas carecem de preparo prévio das amostras a serem analisadas, pois as informações químicas da composição da cadeia graxa são obtidas da derivatização dos triacilgliceróis, aos quais possuem alto peso molecular, o que impossibilitaria a análise da cromatografia gasosa. Essas reações demandam tempo e custo adicional para determinar a composição química dos óleos em geral.

Quadro 2: Técnicas aplicadas a análise de óleos e gorduras vegetais

Amostra	Método Analítico	Aplicação	Referência
Azeite de oliva	FT-NIR	Parâmetro sensorial para PDO	SINELLI <i>et al.</i> , 2010
	RMN	Autenticação	ALONSO-SALCES, 2011
	CG-EM	Avaliação da mistura do azeite de oliva com outros óleos	MONFREDA, M. e GOBBI L., 2012
	CG-DIC	Discriminação geográfica	PIZARRO, 2011
	RAMAN	Cultivo e autenticação	KORIFI <i>et al.</i> , 2011
Óleo de camélia	RMN	Adulteração	SHI <i>et al.</i> , 2018
Óleo de amendoim	RMN	Adulteração	ZHU <i>et al.</i> , 2017
Óleo de andiroba	FTIR-HATR	Controle de qualidade	CABRAL <i>et al.</i> , 2013
	EM	Adulteração	DE SANTANA <i>et al.</i> , 2018
Óleo de palma	CLAE-UV	Origem geográfica	OBISESAN <i>et al.</i> , 2017

Dependendo do objetivo do trabalho, geralmente são utilizadas separações cromatográficas líquidas hifenadas a outros instrumentos, como o ultravioleta visível (UV-Vis), detector de arranjo diodos (DAD) e espectrômetro de massas (EM) (BAMBA e FUKUSAKI, 2006; VERPOORTE *et al.*, 2008; TYAGI *et al.*, 2010). Meziane-Assami e colaboradores (2013) analisaram o óleo essencial de alecrim (*Rosmarinus officinalis* L.) para caracterizar compostos alvo e associar a sua região geográfica; os compostos fenólicos foram analisados por CLAE-DAD e através do CG-EM identificaram de 88 a 99% dos terpenos do óleo essencial. Outra técnica que também está sendo utilizada em óleos é espectroscopia Raman. Através dela, em estudos já publicados, pôde-se avaliar o óleo essencial extraído das folhas, galhos e madeira da árvore do pau-rosa (*Aniba rosaeodora* Ducke), caracterizando os constituintes do óleo com o intuito de garantir a origem sustentável do óleo, uma vez que o mesmo é uma valiosa fonte econômica de linalol para as indústrias de perfume e a exploração pela madeira é proibida (ALMEIDA, 2013).

Cabral e colaboradores (2013) realizaram um trabalho de *fingerprint* com o óleo de andiroba por meio da técnica de espectrometria de massas aliado a técnicas quimiométricas, cujo objetivo foi identificar o tipo de óleo de andiroba com relação a sua natureza (bruto ou refinado) e controle de qualidade (adulteração por óleo de soja). Através dos dados espectrométricos foi possível realizar o estudo quimiométrico por PCA que distinguiu a andiroba adulterada do óleo de andiroba padrão e também o óleo de soja, o que permitiu identificar esses comportamentos foi a composição da cadeia graxa dos distintos óleos, bem como a presença dos limonoides presentes apenas no óleo de andiroba. Outra técnica que também foi utilizada para o estudo de adulteração do óleo de andiroba foi a espectrometria de FTIR-HATR aliada a métodos quimiométricos de classificação, como o I-PLS. Neste trabalho foram utilizados tanto o óleo de soja quanto o óleo de milho para realizar as adulterações do óleo de andiroba, em que foi possível determinar com precisão de 94-100% a classificação de amostras adulteradas e autênticas de andiroba (DE SANTANA *et al.*, 2018).

### 2.2.2 RMN aplicada à adulteração de óleos

A espectroscopia de Ressonância Magnética Nuclear (RMN) vem crescendo em estudo de autenticação de amostras vegetais e em alimentos, principalmente com experimentos de RMN de hidrogênio ( $^1\text{H}$ ). A RMN é uma medida física das ressonâncias dos núcleos atômicos em um forte campo magnético, que permite a caracterização e identificação das substâncias com propriedades químicas diferentes. Os espectros de RMN são de característica ímpar para cada substância, o que possibilita a construção de um banco de dados, cujo somatório das informações obtidas das várias substâncias encontradas numa amostra geram sua impressão digital. Além de ser uma técnica reproduzível, outra vantagem é que no procedimento para análise em RMN, suas limitações são apenas com relação à escolha do solvente e a intensidade do campo magnético (VERPOORTE *et al.*, 2007; LAMANNA, 2013; GOWDA e RAFTERY, 2014).

A reprodutibilidade da técnica de RMN, contanto que seja medido nas mesmas condições, está correlacionada a três informações: o deslocamento químico dos núcleos, as constantes de acoplamento e a integral de cada sinal. Os dados espectrais de RMN não se alteram quando se mantém a intensidade do campo aplicado, solvente, pH e temperatura (KOOY *et al.*, 2009). O que não ocorre quando se empregam métodos cromatográficos e de espectrometria de massa (VERPOORTE *et al.*, 2007; KOOY *et al.*, 2009).

Espectrômetros de RMN realizam experimentos unidimensional (1D) e bidimensional (2D) para distintos núcleos atômicos, tais como:  $^1\text{H}$ ,  $^2\text{H}$ ,  $^7\text{Li}$ ,  $^{11}\text{B}$ ,  $^{13}\text{C}$ ,  $^{14}\text{N}$ ,  $^{19}\text{F}$  e  $^{31}\text{P}$  (SINGH e BLUMICH, 2016). Em estudos por quimiometria, geralmente são utilizados experimentos com núcleos atômicos de hidrogênio ( $^1\text{H}$ ) e/ ou carbono ( $^{13}\text{C}$ ), devido à presença em várias moléculas orgânicas. Contudo, a abundância do isótopo é fundamental para que os experimentos sejam obtidos com menor quantidade de material e de tempo, assim os núcleos de hidrogênio são os mais investigados cuja abundância natural é de 99,98% (GOWDA e RAFTERY, 2014).

A técnica de RMN de  $^1\text{H}$  vem sendo bastante empregada em diversos estudos qualitativos e quantitativos com óleos e gorduras, bem como, caracterização e quantificação de óleos vegetais, determinação da origem geográfica, identificação de adulteração, controle de qualidade de óleos comestíveis, controle reacional da produção de biodiesel, otimização de processos, teor de lipídeos em carnes e

embutidos, dentre outros. Vários trabalhos envolvem principalmente estudos que tratam do controle de qualidade e identificação de adulteração do óleo de oliva, pois o valor agregado a esse produto é bastante valorizado, o que estimula a grande comercialização local e mundial. Através dos experimentos de RMN de  $^1\text{H}$  é possível verificar se o óleo de oliva foi adulterado com óleos mais baratos e que possuam características químicas próximas, pois através do espectro é possível identificar substâncias que não pertencem ao óleo. Caso o óleo utilizado na adulteração possuir a mesma composição da cadeia graxa, será possível quantificar e identificar a substância que ultrapassa a concentração que não é normalmente encontrada no óleo de oliva, para isso é necessário fazer um estudo quantitativo ou utilizar a quimiometria para auxiliar na classificação dos óleos investigados (ABBAS e BAETEN, 2016; GÓMEZ-CARAVACA *et al.*, 2016).

### 2.3 Quimiometria e algumas fundamentações

A quimiometria surgiu no período em que houve um grande avanço computacional na área da pesquisa, onde as técnicas instrumentais começaram a gerar uma gama de informações que levavam muito tempo para serem interpretáveis. Logo, a quimiometria veio auxiliar os químicos a darem explicações de suas atividades experimentais tanto depois da obtenção dos dados quanto antes de adquiri-los (BRERETON, 2014). O termo quimiometria foi denominado pela primeira vez por Svant Wold em 1971 que juntamente com Kowalski definiram a quimiometria como: “A arte de extrair informações quimicamente relevantes dos dados produzidos em experimentos químicos” (WOLD, 1995). A quimiometria auxilia na interpretação das informações pertinentes ao estudo, avaliando as tendências e o comportamento dos dados, principalmente quando se trata de uma abordagem geral dos metabólitos presentes numa amostra, denominado de estratégia não direcionada (*non-targeted*) (CUADROS-RODRÍGUEZ *et al.*, 2016).

Os métodos quimiométricos mais utilizados pela comunidade científica são os multivariados, sempre associados a um grande volume de dados a serem interpretados. Dependendo do objetivo do trabalho, os métodos quimiométricos podem ser utilizados para fins diversos, como por exemplo, o “reconhecimento de padrões e classificação” de alimentos. Este tipo de estudo promove o controle de qualidade de produtos cuja qualidade está associada às suas características químicas



e que podem estar correlacionadas ao local de origem e/ ou produção local do produto. Geralmente esse método de reconhecimento de padrões avalia a semelhança ou diferença de vários tipos de amostras e podem ser subdivididos em métodos supervisionados e não supervisionados (FERREIRA, 2015).

Os métodos não supervisionados não necessitam de conhecimento prévio da matriz da amostra, tanto que também podem ser denominados como uma análise exploratória dos dados. Eles avaliam as tendências que os dados apresentam, quanto a sua discriminação ou similaridade entre si. Os principais e mais utilizados métodos não supervisionados são Análise de Componentes Principais (PCA - *Principal Component Analysis*) e Análise Hierárquica de grupos (HCA - *Hierarchical Cluster Analysis*). Já os métodos supervisionados geralmente possuem alguma informação que pode auxiliar na classificação dos dados que estão sendo analisados e, essa informação é indispensável para construir um modelo de classificação. Dentre os métodos mais utilizados, encontra-se a análise dos Quadrados Mínimos Parciais (I-PLS - *Partial Least Square*) e suas variações; bem como o SIMCA - *Soft Independent Modeling of Class Analogy* (FERREIRA, 2015).

A Análise de Componentes Principais (PCA) foi introduzida por Karl Pearson em 1901, cuja proposta foi reduzir a dimensionalidade do banco de dados antes de promover a análise do comportamento e as tendências dos dados. Como a organização dos dados é em formato de matriz ( $X_{ij}$ ), as amostras (i) são as linhas e as variáveis (j) são as colunas. A combinação linear das variáveis originais reduz a quantidade de variáveis a serem discutidas, gerando um novo grupo de variáveis chamadas de componentes principais. A componente principal 1 é aquela que sempre carrega consigo a maior quantidade de informações, pois ela se move na direção que descreve a maior variância dos dados. Apesar da primeira componente descrever a maior quantidade informação dos dados, se a natureza da matriz for muito complexa, necessitará de mais de uma componente principal para explicar (SENA e POPPI, 2000; FERREIRA, 2015).

Os Quadrados Mínimos Parciais (I-PLS), denominado por Herman em 1970, trabalha de forma semelhante à PCA reduzindo a dimensionalidade dos dados. A diferença entre eles é que o I-PLS é uma análise supervisionada; ou seja, além de possuir a matriz X que detém as variáveis latentes (conhecidas como componentes principais em PCA), também possui uma matriz Y, chamada de matriz resposta. Assim, a matriz X aponta na direção de maior variância da matriz resposta (matriz Y),

cuja mesma contém informações conhecidas que podem ser utilizadas para classificar as amostras da matriz X (FERREIRA, 2015).

A Análise de Agrupamento Hierárquico reúne amostras (ou objetos) que possuem alguma similaridade entre si, formando agrupamentos que são dispostos em um dendrograma. O gráfico do dendrograma pode ser interpretado da seguinte forma, quanto menor a distância entre os objetos maior será a similaridade entre eles; o contrário também é verdadeiro, quanto maior a distância, menor a similaridade entre eles (KAUFMAN e ROUSSEEUW, 2005).

Outra análise de cluster bastante utilizada na ecologia e em problemas geoestatísticos é o *Fuzzy C-Means* (FCM), que também pode ser utilizado em dados quimiométricos. A diferença entre os dois métodos é que a HCA é uma abordagem clássica denominada de “*crisp*”, a qual determina que uma amostra pertence unicamente a um determinado grupo; já o FCM é uma análise nebulosa que considera que uma amostra pode ter níveis ou graus de pertencimento em um ou mais grupos. A representação gráfica do FCM é semelhante à PCA, em que apresenta a formação dos grupos e o grau de pertencimento que a amostra pode ter em relação a outro grupo (BEZDEK, EHRLICH e FULL, 1984).

### 3. OBJETIVOS

#### 3.1 Objetivo geral

Empregar a técnica de Ressonância Magnética Nuclear (RMN) aliado à quimiometria com o intuito de avaliar o óleo de andiroba e sua possível adulteração.

#### 3.2 Objetivos específicos

- ✓ Obter espectros unidimensionais de RMN de  $^1\text{H}$  dos óleos de andiroba e soja e misturas em diferentes proporções para as análises de adulteração por quimiometria;
- ✓ Obter os espectros bidimensionais para realizar a caracterização dos óleos, tais como: COSY, HMBC E HSQC;
- ✓ Realizar a análise multivariada não-supervisionada dos dados obtidos dos espectros de RMN de  $^1\text{H}$ , tais como: Análise de Componentes Principais (PCA), Análise de Agrupamento Hierárquico (HCA) e *Fuzzy C-Means* (FCM);
- ✓ Realizar a análise multivariada supervisionada dos dados obtidos dos espectros de RMN de  $^1\text{H}$  por meio do modelo de regressão por mínimos quadrados parciais, I-PLS (*Partial Least Squares*);
- ✓ Identificar os sinais que possam indicar a adulteração de óleo de andiroba por óleo de Soja;

## 4. METODOLOGIA

### 4.1 Obtenção dos óleos de andiroba

As amostras utilizadas foram divididas em duas denominações: padrão de óleo de andiroba (óleo de cooperativa e extrato de sementes de andiroba) e óleo de andiroba comercial (adquiridas em comércios). As amostras padrão de óleo de andiroba são de cooperativas do município de Manaquiri coletadas em 2012 e os extratos das sementes já se encontravam no banco de amostras do grupo de pesquisa. Segue a tabela com a descrição sucinta das amostras e seus devidos códigos.

Quadro 3: Amostras de óleo de andiroba utilizadas nos experimentos de RMN.

Código	Descrição	Localidade
A01	Padrão de óleo de andiroba 1, óleo da cooperativa A	Manaquiri/ AM
A02	Padrão de óleo de andiroba 2, óleo da cooperativa B	
A03	Padrão de óleo de andiroba 3, óleo da cooperativa C	
A04	Padrão de óleo de andiroba 4, óleo da cooperativa D	
A06	Padrão de óleo de andiroba 5, extrato hexânico A	
A16	Padrão de óleo de andiroba 6, extrato hexânico B	
A17	Padrão de óleo de andiroba 7, extrato hexânico C	
A05	Óleo de andiroba comercial 1	Manaus
A09	Óleo de andiroba comercial 2	
A10	Óleo de andiroba comercial 3	
A11	Óleo de andiroba comercial 4	
A12	Óleo de andiroba comercial 5	
A14	Óleo de andiroba comercial 6	
A15	Óleo de andiroba comercial 7	
A22	Óleo de andiroba comercial 8	
A24	Óleo de andiroba comercial 9	
A23	Óleo de andiroba comercial 10	
A25	Óleo de andiroba comercial 11	
A26	Óleo de andiroba comercial 12	

As exsiccatas da *C. guianensis* coletadas pelo grupo, encontram-se depositadas no Herbário do INPA, sob os números 178.658 e 192.616, referente à amostra A06. As amostras comerciais foram adquiridas nas feiras e farmácias locais

da cidade de Manaus-AM e acondicionadas à temperatura ambiente. Logo, ao total foram utilizadas 19 amostras de óleo de andiroba neste estudo.

#### 4.2 Materiais e equipamentos

- ✓ Clorofórmio deuterado ( $\text{CDCl}_3$ ) 99,80% D da Cambridge Isotope Laboratories, Inc.
- ✓ Tetrametilsilano (TMS) 99,50% D da SIGMA/ALDRICH
- ✓ Tubo de RMN 500 MHz de 5 mm da NORELL
- ✓ Capilar com solução de 0,50% de TMS em  $\text{CDCl}_3$  99,60% D da Adrich
- ✓ Pipeta de 200  $\mu\text{L}$  e 1000  $\mu\text{L}$  da *Eppendorf Research*
- ✓ Ponteiras de 200  $\mu\text{L}$  e 1000  $\mu\text{L}$  da *Eppendorf Research*
- ✓ Pipeta *multipette plus Eppendorf Research*
- ✓ Espectrômetro BRUKER, modelo FOURIER 300, magneto 300 SB Ultra Shield TM, 7,05 T (300 MHz para  $^1\text{H}$  e 75 MHz para  $^{13}\text{C}$ ), sonda de 5 mm com dois canais (DUAL  $^{13}\text{C}$ - $^1\text{H}$ / D), com detecção direta e gradiente de campo no eixo Z; instalado na Central Analítica do Laboratório Temático de Química de Produtos Naturais (CA-LTQPN) do INPA.

#### 4.3 Obtenção dos espectros bidimensionais para caracterização dos óleos

As amostras de óleo de andiroba padrão e de soja foram preparadas em um tubo de RMN no qual foi transferido um volume de 50  $\mu\text{L}$  de óleo e 500  $\mu\text{L}$  de solução de  $\text{CDCl}_3$  com 0,50% de TMS v/v, altura do volume total da solução permaneceu na região da bobina, conforme visualizado no medidor de profundidade do equipamento.

Para a caracterização dos compostos foram utilizados experimentos bidimensionais com gradiente de campo (COSY, HSQC e HMBC), empregando respectivamente as seguintes sequências de pulsos: *cosygpqf*, *hsqcgpph* e *hmbcgpplndqf* (denominação Bruker).

Todos os experimentos bidimensionais COSY, HSQC e HMBC foram realizados com largura espectral (SWH) de 20,33 ppm em F2; e de 20,33 ppm, 189,99 ppm e 219,99 ppm, respectivamente em F1.

Os parâmetros para os experimentos bidimensionais COSY, HSQC, HMBC foram os mesmos: tempo de espera entre cada aquisição de 1 s (d1), com 2048 pontos durante a aquisição em F2 e 256 em F1.

#### 4.4 Preparo das amostras do Banco de dados 1 (andiroba padrão, andiroba comercial e soja) para análise quimiométrica não supervisionada

Ao **banco de dados 1** foram adicionadas duas amostras de óleo de soja comercial (S02 e S03) para avaliar o comportamento das amostras de andiroba frente a esta matriz que é bastante utilizada na adulteração de óleos vegetais. Portanto, o total de amostras para o **banco de dados 1** foram de 21 amostras no total.

Quadro 4: Amostras do **Banco de dados 1** para análise quimiométrica não supervisionada

Código	Descrição	Localidade
A01	Padrão de óleo de andiroba 1, óleo da cooperativa A	Manaquiri/ AM
A02	Padrão de óleo de andiroba 2, óleo da cooperativa B	
A03	Padrão de óleo de andiroba 3, óleo da cooperativa C	
A04	Padrão de óleo de andiroba 4, óleo da cooperativa D	
A06	Padrão de óleo de andiroba 5, extrato hexânico A	
A16	Padrão de óleo de andiroba 6, extrato hexânico B	
A17	Padrão de óleo de andiroba 7, extrato hexânico C	
A05	Óleo de andiroba comercial 1	Manaus
A09	Óleo de andiroba comercial 2	
A10	Óleo de andiroba comercial 3	
A11	Óleo de andiroba comercial 4	
A12	Óleo de andiroba comercial 5	
A14	Óleo de andiroba comercial 6	
A15	Óleo de andiroba comercial 7	
A22	Óleo de andiroba comercial 8	
A24	Óleo de andiroba comercial 9	
A23	Óleo de andiroba comercial 10	
A25	Óleo de andiroba comercial 11	
A26	Óleo de andiroba comercial 12	
S02	Óleo de SOJA comercial 2	Manaus
S03	Óleo de SOJA comercial 3	

O modo de preparo das amostras foi diretamente realizado ao tubo de RMN, ao qual foi transferido um volume de 50  $\mu\text{L}$  de óleo e 500  $\mu\text{L}$  de  $\text{CDCl}_3$ . Em seguida, colocava-se o capilar que continha uma solução de 0,50% v/v de TMS; a altura do volume total da solução permaneceu na região da bobina, conforme visualizado no

medidor de profundidade do equipamento. Como o padrão de referência é bastante volátil e todas as amostras eram preparadas em  $\text{CDCl}_3$ , optou-se em preparar uma solução de 0,50% v/v de TMS em  $\text{CDCl}_3$  e depois foi transferida para um capilar de vidro, que posteriormente foi selado com o intuito de torná-lo hermético. Todas as medidas foram realizadas em triplicatas. Como todos os espectros adquiridos foram realizados em triplicatas, para o banco de dados 1 foram obtidos o total de 63 espectros de RMN de  $^1\text{H}$  unidimensional.

#### 4.5 Preparo das amostras do Banco de dados 2 para análise quimiométrica supervisionada

O **Banco de dados 2** é formado pelas amostras da matriz de supervisão (tabela 1, pg.38) mais as amostras da matriz de investigação ou de concentração desconhecida (quadro 5, pg. 39). Como todos os espectros adquiridos foram realizados em triplicatas, para o banco de dados 2 foram obtidos o total de 93 espectros de RMN de  $^1\text{H}$  unidimensional.

Tabela 1: Matriz de supervisão ou matriz X apresenta o grau adulteração do óleo de andiroba com óleo de soja

Mistura	Andiroba, $\mu\text{L}$	Soja, $\mu\text{L}$	Adulteração, %
A01	50,0	-	0
B01	47,5	2,5	5
B02	45,0	5,0	10
B03	42,5	7,5	15
B04	40,0	10,0	20
B05	37,5	12,5	25
B06	35,0	15,0	30
B07	32,5	17,5	35
B08	30,0	20,0	40
B09	27,5	22,5	45
B10	25,0	25,0	50
B11	22,5	27,5	55
B12	20,0	30,0	60
B13	17,5	32,5	65
B14	15,0	35,0	70
B15	12,5	37,5	75
B16	10,0	40,0	80
B17	7,5	42,5	85
B18	5,0	45,0	90
B19	2,5	47,5	95
S01	-	50,0	100

1 - O óleo padrão de andiroba (A01) foi o controle negativo de adulteração (0% de soja); 2- O óleo de soja (S01) foi o valor extremo de 100% de adulteração.

Quadro 5: Matriz de investigação ou matriz Y apresenta as amostras de concentração desconhecida de possível adulteração com óleo de soja

Código	Descrição	Localidade
A05	Comercial 1	Manaus/ AM
A09	Comercial 2	
A10	Comercial 3	
A12	Comercial 5	
A15	Comercial 7	
A22	Comercial 8	
A24	Comercial 9	
A23	Comercial 10	
A25	Comercial 11	
A26	Comercial 12	

As 19 misturas binárias de óleo de andiroba com óleo de soja<sup>2</sup> foram preparadas variando em uma escala de 5% de adulteração (5% a 95% v/v). Todas as misturas foram preparadas da mesma forma que as amostras da análise não supervisionada, ou seja, diretamente no tubo de RMN e solubilizadas com CDCl<sub>3</sub>, em seguida foi adicionado o tubo capilar com o padrão de TMS.

#### 4.6 Aquisição dos espectros de RMN de <sup>1</sup>H para análise quimiométrica

Todos os espectros obtidos para análise quimiométrica utilizaram o procedimento descrito a seguir. A amostra foi inserida no espectrômetro e esperou-se 10 min para que permanecesse em equilíbrio térmico com a sonda. Como o solvente utilizado foi o CDCl<sub>3</sub>, realizou-se o bloqueio do campo (*lock*) com o deutério do solvente, a homogeneização do campo foi realizada manualmente, até o sinal do TMS atingir a meia altura de aproximadamente 0,6 Hz/ppm. O equipamento já vem com ajuste automático da afinação e correspondência da sonda (*tune e match*) e não tem como ajustar manualmente. Todos os espectros foram obtidos utilizando a sequência de pulso ZG (denominação da Bruker). Os espectros foram adquiridos com 64 promediações (NS), largura espectral de 20,33 ppm, tempo de aquisição entre cada varredura 5,3687091 s (AQ) com 65.536 pontos durante a aquisição (TD), a frequência do centro do espectro (O1) foi de 5 ppm. Para cada amostra, foi realizada a calibração do pulso de 90° (P1), a estimativa do tempo de espera entre cada aquisição e o

<sup>2</sup> S01 – óleo de soja comercial mais popular.



próximo pulso (d1), também o tempo de relaxação longitudinal (T1). Para a aquisição desses dados foram utilizados 65,536 pontos (SI).

#### **4.7 Análise quimiométrica dos Banco de dados 1 e Banco de dados 2**

Para a realização do estudo quimiométrico do óleo de andiroba foram utilizados os programas: TopSpin 3.5pl7 (disponibilizada pela Bruker para acadêmicos), o MATLAB versão trial, o R-3.5.2 e o RStudio 1.1.463 para Windows, todos disponíveis em versões gratuitas.

##### **4.7.1 Pré-tratamento dos dados no programa TopSpin de RMN de $^1\text{H}$ para análise quimiométrica**

A transformação dos dados foi aplicada igualmente a todas as amostras, onde todos os espectros foram processados através do programa TopSpin 3.5pl7. Em que foram aplicados filtros de transformada de Fourier (FT), o ajuste da fase (APK) e a correção da linha de base (ABSN) foi realizada automaticamente e sem integração dos sinais; também não foi aplicada a multiplicação exponencial ( $I_b = 0,0$  Hz). A calibração dos espectros foi definida pelo sinal do TMS em 0 ppm.

##### **4.7.2 Montagem da Matriz, pré-tratamento e pré-processamento dos dados no MATLAB e RStudio**

A exportação dos dados espectroscópicos foi realizada pelo algoritmo “convbin2csv” adaptado ao programa TopSpin e o tratamento dos dados foram executados no programa RStudio, através das rotinas disponibilizadas pelo professor Dr. Luiz Henrique Keng Queiroz Júnior, do Laboratório de Ressonância Magnética Nuclear LABRMN, da Universidade Federal de Goiás (UFG).

O algoritmo “convbin2csv” exporta os dados do espectro em um arquivo com extensão .CSV (*Comma Separated Values*), gerando uma planilha em que o deslocamento químico e intensidade são separados por vírgula numa mesma coluna. Cada linha representa os distintos deslocamentos e intensidades de um mesmo espectro. Após a exportação de todos os espectros para o formato CSV foi realizado o alinhamento dos espectros com o algoritmo ICOSHIFT no programa MATLAB versão trial, cujo objetivo é de ajustar a diferença de deslocamentos químicos de mesmos sinais de espectros diferentes. Foi determinada uma faixa de trabalho com a frequência mínima e a frequência máxima do espectro (-1,00 a 8,00 ppm),

acompanhar figura 6. Posteriormente, no programa RStudio, procedeu-se com a limpeza das regiões ruidosas do espectro e de sinais que não foram utilizados na quimiometria; como por exemplo, o sinal do solvente ( $\text{CDCl}_3$ ) e o grupo metileno de cadeias alifáticas ( $-\text{CH}_2-$ ). Essa limpeza é denominada regiões de exclusão, no qual para a construção dos dois bancos de dados foram excluídas 9 regiões, conforme a figura 7. Os intervalos de exclusão podem ser acompanhados na tabela 2.

Figura 6: Região dos hidrogênios beta da unidade do glicerol antes e após o alinhamento no programa MATLAB através do algoritmo ICOSHIFT.

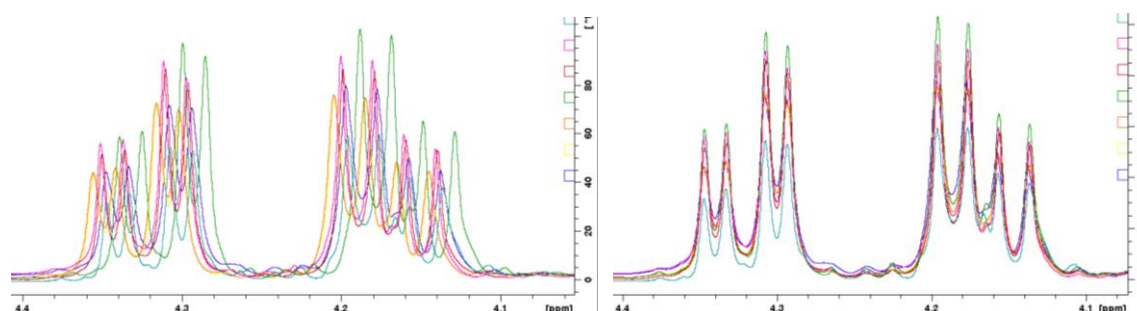
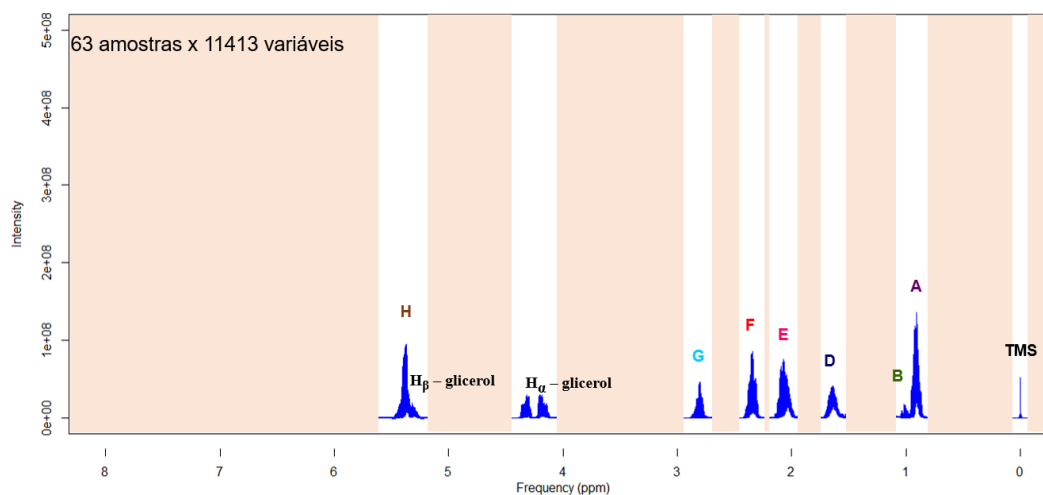


Tabela 2: Intervalos de frequência das regiões de exclusão dos espectros de RMN de  $^1\text{H}$  dos óleos vegetais.

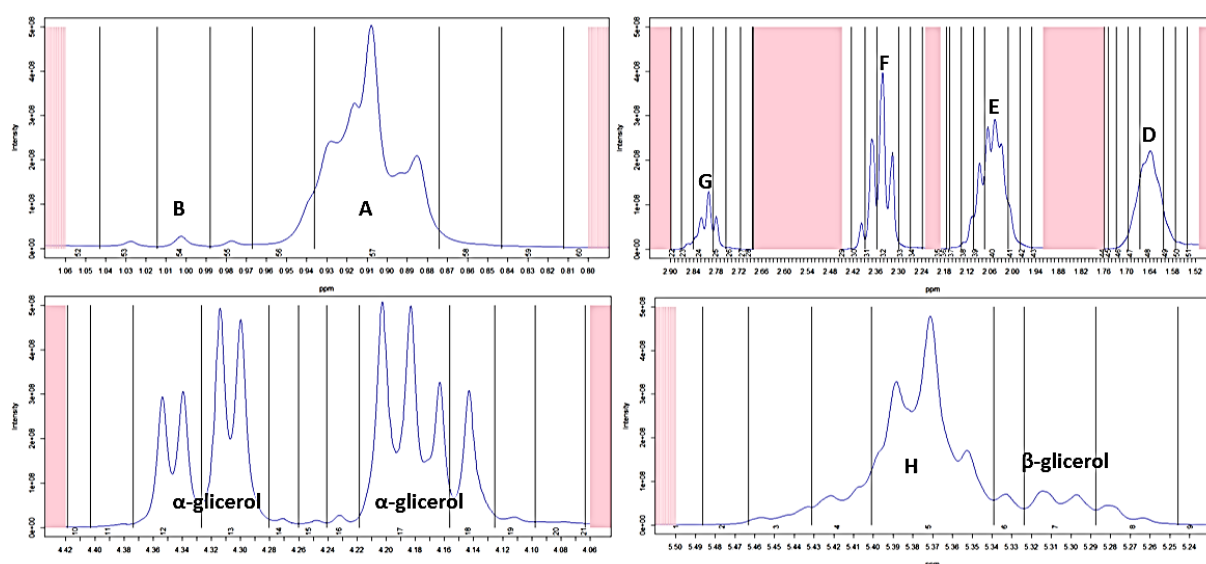
Nº de exclusão	Região do espectro	Intervalo da região, ppm
1	Ruído	-1,00 a -0,05
2		0,04 a 0,80
3	Metileno de cadeias alifáticas	1,06 a 1,51
4		1,76 a 1,92
5		2,19 a 2,23
6	Ruído	2,45 a 2,68
7		2,90 a 4,06
8		4,42 a 5,20
9	Ruído e $\text{CDCl}_3$	5,50 a 8,00

Figura 7: Espectro de RMN de  $^1\text{H}$  dos óleos vegetais e as regiões de exclusão representadas pela cor rosa



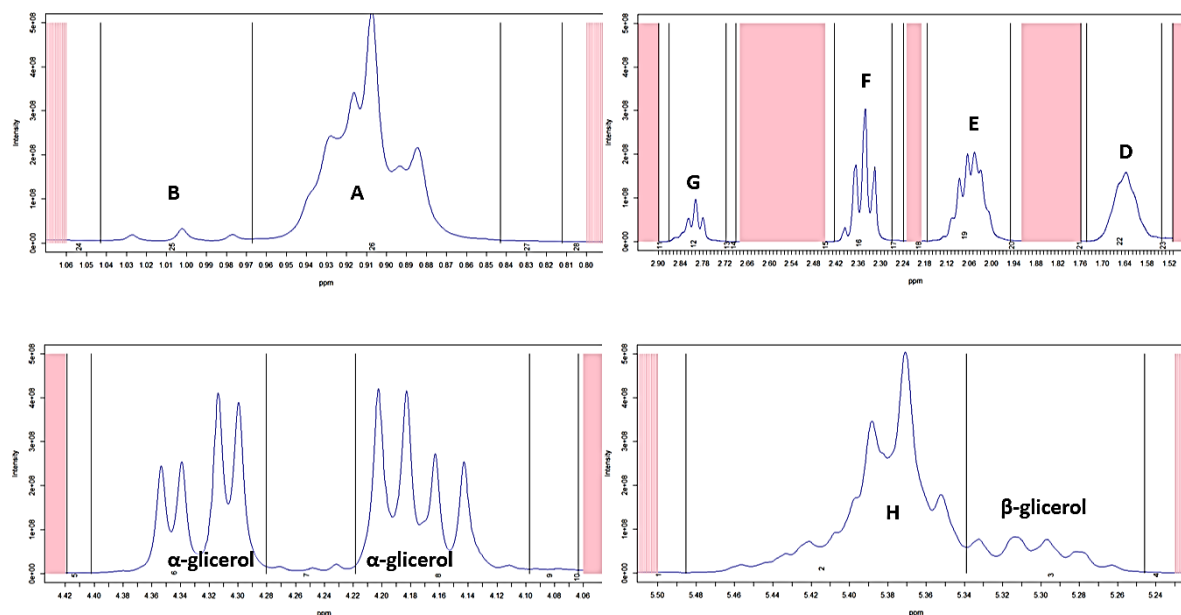
Inicialmente, cada espectro do **Banco de dados 1 e 2** possuía 65,536 pontos ( $f_{\min} = -5,11$  ppm e  $f_{\max} = 15,20$  ppm), após determinar a faixa de trabalho e realizar as exclusões, cada espectro ficou com 6.407 pontos. Posteriormente, para os dados multivariados foram determinadas as áreas dos sinais, uma vez que os dados exportados não estavam com as áreas integradas. Portanto, para determinar a área de um sinal para análise quimiométrica foram criados os *buckets* ou *binning* automático, cujo tamanho dos intervalos foram de 0,03 ppm com grau de liberdade de 30%, gerando 64 *buckets* (figura 8).

Figura 8: Representação dos *buckets* automáticos cortando as áreas de sinais do espectro de RMN de  $^1\text{H}$  dos óleos vegetais



Ao plotar o espectro com os *buckets* automáticos (Figura 8), verificou-se que havia cortes indevidos de alguns sinais, o que poderia interferir na integração da área desses sinais. Portanto, houve a necessidade de realizar a otimização desses *buckets* manualmente, conferindo agora um total de 31 *buckets* (Figura 9).

Figura 9: Representação dos *buckets* após a otimização manual



Com o tamanho dos *buckets* ajustados, foi possível normalizar os espectros pela área do TMS com o objetivo de ajustar os espectros devido à variação de temperatura. Como o TMS tinha a concentração definida, pois encontrava-se em um tubo capilar hermético, utilizou-se a área desse sinal para realizar a normalização de todas as amostras.

O escalamento utilizado no pré-processamento das variáveis foi o escalamento por Pareto. Com o pré-tratamento e pré-processamento realizados, pôde-se prosseguir para a análise multivariada exploratória e classificatória. É importante salientar que todos os espectros foram pré-processados e pré-tratados igualmente.

Realizaram-se três análises exploratórias: a HCA foi utilizado o pacote “ChemoSpec”, cujo método foi o “Ward”. A análise *Fuzzy C-Means* (FCM) foi realizada com o pacote “cluster” através da distância Euclidiana e a PCA robusta com o pacote “pcaPP”, todos disponíveis para o programa R. Já a análise multivariada classificatória por I-PLS foi carregado o pacote PLSR também disponível ao programa R.

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Caracterização química dos óleos vegetais

#### 5.1.1 Atribuição dos sinais nos espectros de RMN de $^1\text{H}$ unidimensional dos óleos de andiroba e soja.

Os espectros de óleos vegetais possuem um perfil espectral muito semelhante, pois a composição química é de 95-98% de triacilgliceróis, que são unidades de ácidos graxos esterificados a unidades de glicerol. Devido à similaridade química dos diferentes ácidos graxos, os sinais de ressonância magnética ressonam juntos e apresentam sobreposição de sinais, em que o espectro de RMN de  $^1\text{H}$  de óleos vegetais apresenta um pequeno número de picos característicos (Tabela 4). Logo, o que caracteriza e diferencia um óleo vegetal de outro são as proporções dos ácidos graxos presentes na sua forma livre (AGL) ou esterificada (ésteres glicéridos). Os triacilgliceróis encontrados nos óleos de andiroba e soja foram apresentados na tabela 3 que discrimina as proporções e os tipos de ácidos graxos pela técnica de cromatografia gasosa.

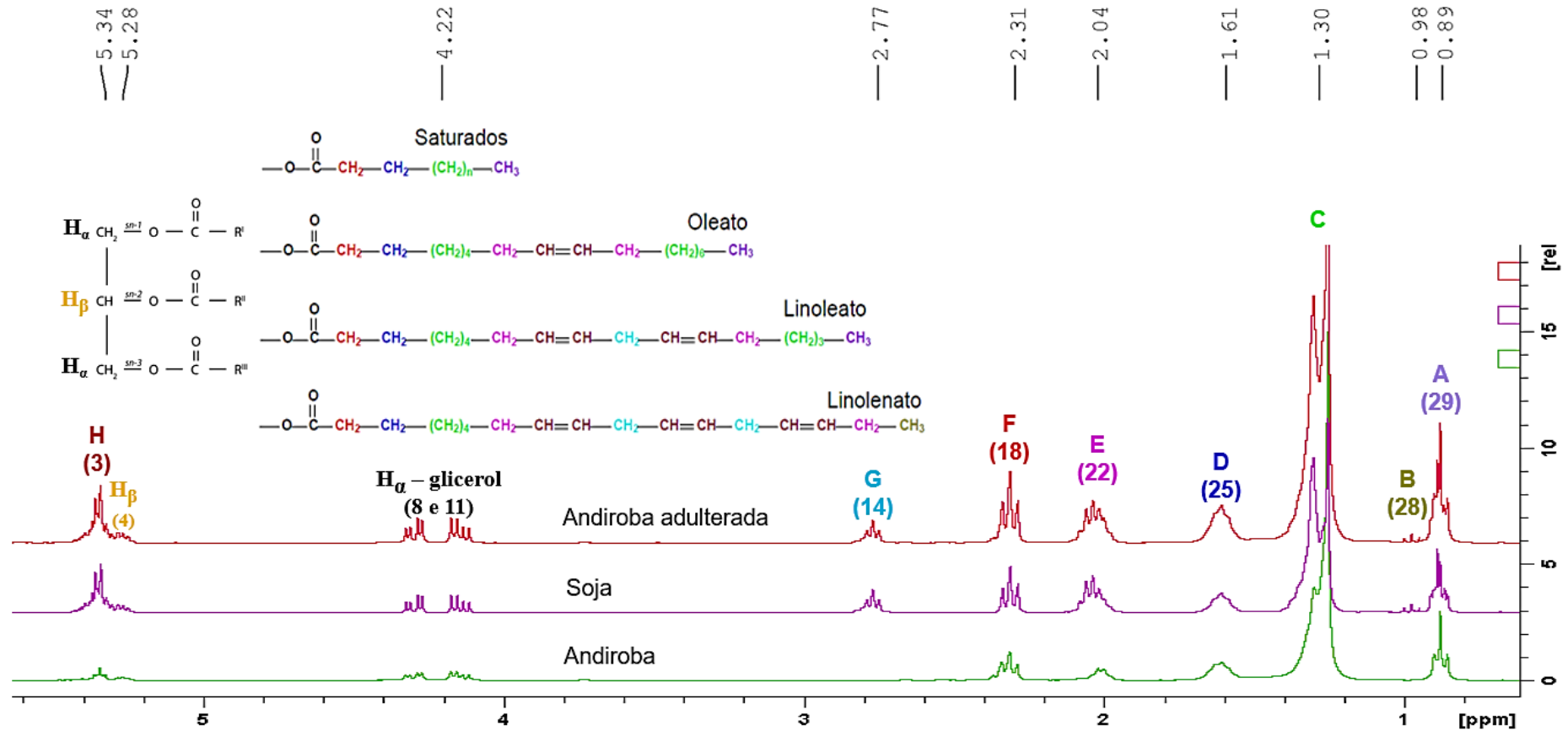
Tabela 3: Composição da cadeia graxa dos óleos de andiroba e soja por cromatografia gasosa (CG)

Ácido graxo	Cadeia	% de Área	
		Andiroba <sup>1</sup>	Soja <sup>2</sup>
Palmítico	C 16:0	25,76	12,00
Esteárico	C 18:0	9,08	4,00
Oleico	C 18:1 <i>cis</i> -9	51,81	24,00
Linoleico	C 18:2 <i>cis</i> -9,12	8,30	52,00
Linolênico	C 18:3 <i>cis</i> -9,12,15	-	8,00

<sup>1</sup>Cabral *et al*, 2013; <sup>2</sup>Dhakal *et al*, 2014.

Os espectros de RMN de  $^1\text{H}$  adquiridos dos óleos vegetais de andiroba e soja, podem ser acompanhados inicialmente na figura 10, na qual, mostrou-se sinais típicos de triacilgliceróis, cuja composição dos ácidos graxos esterificados foi análoga às relatadas em outros trabalhos para diferentes óleos vegetais, inclusive o óleo de soja (BARISON, 2010; CABRAL *et al*, 2013; DHAKAL *et al*, 2014).

Figura 10: Espectros de RMN <sup>1</sup>H destacando os tipos de hidrogênios presentes nos óleos de soja, andiroba padrão e adulterada, em que R são os ésteres característicos das amostras.



Fonte: Adaptado de BARISON, 2010.

Comparando os dados de deslocamentos químicos ( $\delta$ ) com a literatura, o sinal A ( $\delta$  0,86-0,90) corresponde aos grupos metila de cadeias alquila dos ésteres oleato, linoleato e todos os saturados, com exceção ao éster linolenato. Ao sinal B ( $\delta$  0,92-1,02) é característico do éster linolenato, pois este grupo metila é mais desblindado devido à proximidade da insaturação na posição C15 (ou  $\omega$ -3), este sinal é encontrado apenas no óleo de soja. O sinal C ( $\delta$  1,28-1,30) é atribuído aos grupos metileno de cadeias alquilas que estão presentes em todos os ácidos graxos saturados e insaturados. Os sinais F ( $\delta$  2,31-2,33) e D ( $\delta$  1,61-1,63) são grupos metilênicos  $\alpha$  e  $\beta$  carbonílicos de todos os ésteres, respectivamente. O grupo  $\alpha$  carbonílico possui deslocamento em 2,31-2,33 ppm devido ao campo anisotrópico do grupo carbonila que desblinda esses hidrogênios, quanto mais afastado do grupo carbonila, menor o efeito e conseqüentemente menor o deslocamento. O sinal E ( $\delta$  2,03) é conhecido como grupo metilênico  $\alpha$  olefínico presente em todos os ésteres insaturados. O sinal G ( $\delta$  2,77-2,79) corresponde ao hidrogênio de grupos metileno entre olefinas, esse sinal é típico de ésteres linoleato e linolenato que se encontra em maior proporção para o óleo de soja, o óleo de andiroba só possui o linoleato e numa proporção baixíssima. O sinal H ( $\delta$  5,34-5,37) é atribuído aos hidrogênios das insaturações, eles possuem o maior deslocamento químico de todos os sinais devido à desblindagem anisotrópica das ligações duplas. O sinal  $H_\beta$  é o hidrogênio beta da unidade do glicerol, assim como os sinais  $H_\alpha$  são os hidrogênios alfa da unidade do glicerol.

Tabela 4: Atribuição dos deslocamentos químicos dos hidrogênios de triacilgliceróis dos óleos vegetais por RMN de  $^1\text{H}$  unidimensional

Compostos	Deslocamento químico, $\delta$ (ppm)			
	Sinal	Soja <sup>1</sup>	Andiroba <sup>1</sup>	Referência <sup>2</sup>
Todos os ácidos exceto linolênico	A	0,86	0,90	0,87
Linolênico	B	0,92	-	1,02
Todas as cadeias alquilas	C	1,28	1,30	1,30
Todas as cadeias alquilas	D	1,61	1,63	1,62
Todos os AG insaturados	E	2,03	2,03	2,03
Todas as cadeias alquilas	F	2,31	2,33	2,32
Ácidos linoleico e linolênico	G	2,77	2,79	2,77
Triacilglicerol	A	4,22	4,24	4,22
Triacilglicerol	B	5,28	5,29	5,26
Todos os AG insaturados	H	5,34	5,37	5,37

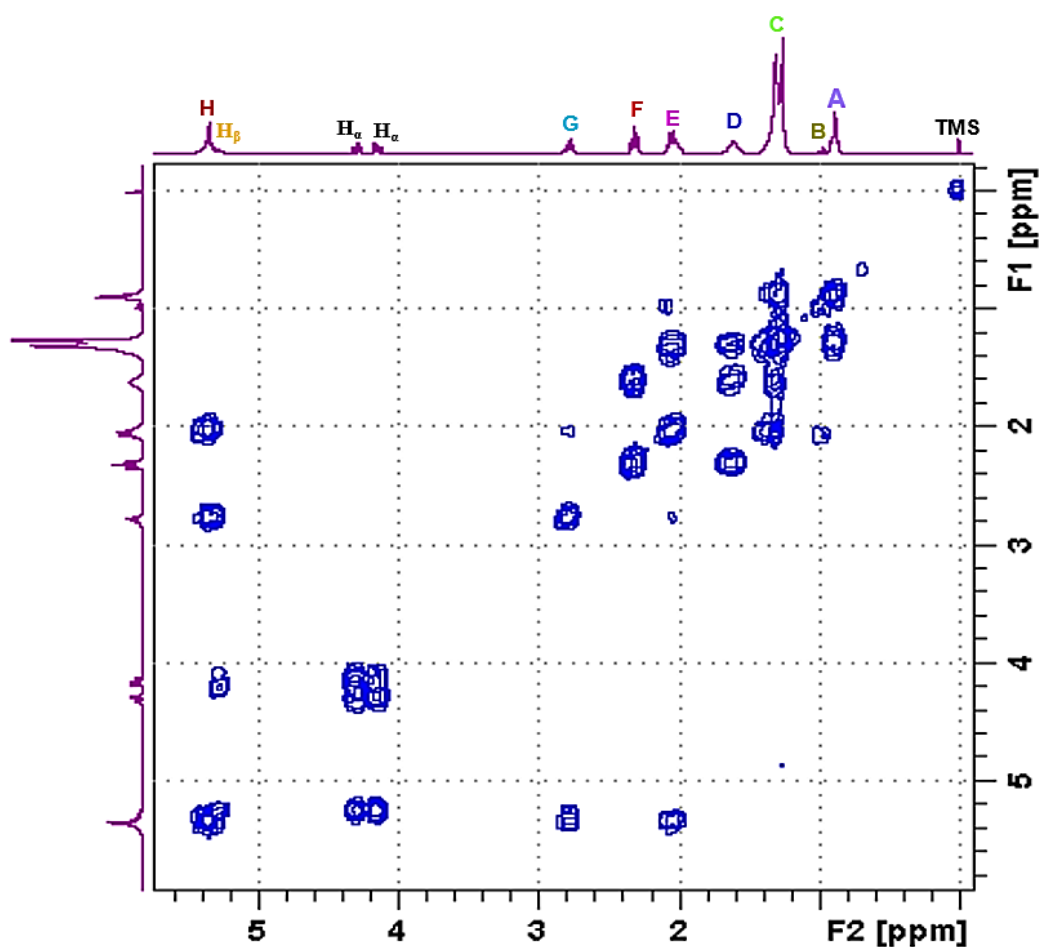
<sup>1</sup>Média da faixa de deslocamentos químicos dos óleos vegetais referenciados pelo TMS em 0,00 ppm.

<sup>2</sup>Popescu *et al*, 2015.

### 5.1.2 Atribuição dos sinais nos espectros de RMN de $^1\text{H}$ bidimensional homonuclear (COSY) dos óleos de soja e andiroba.

O experimento de COSY foi utilizado para avaliar as correlações homonucleares de hidrogênio ( $^1\text{H}$ - $^1\text{H}$ ) e confirmar as atribuições discutidas no experimento de RMN de  $^1\text{H}$  unidimensional aos TAG presente na composição química dos óleos vegetais.

Figura 11: Mapa de contornos COSY adquirido para a amostra de soja (S02)

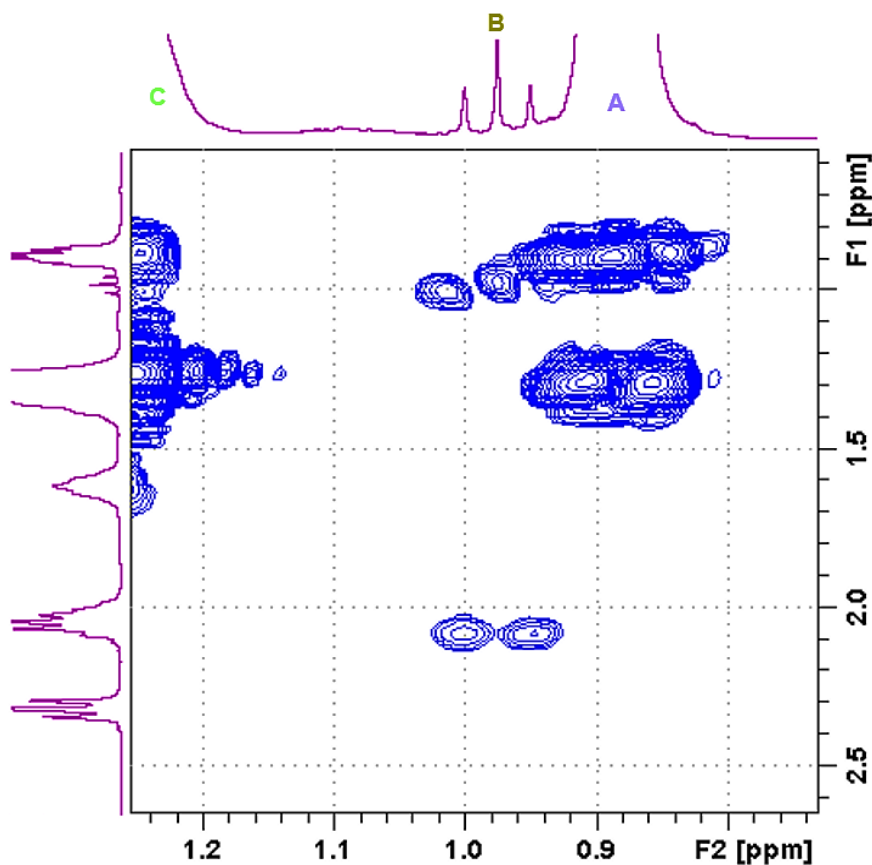


O mapa de contorno COSY (figura 11) do óleo de soja foi utilizado na interpretação do acoplamento dos hidrogênios, pois o mesmo apresenta todos os sinais que estão presentes no óleo de andiroba. A única correlação não existente no óleo de andiroba é o grupo metila de cadeias alquilas do éster linolenato, presente apenas no óleo de soja. O sinal C ( $\delta$ 1,28-1,30) denominado com grupo metileno de cadeias alquilas está presente em todos os ésteres dos óleos de andiroba e soja. Esse sinal se correlaciona com o sinal A ( $\delta$  0,86-0,90) dos grupos metila das cadeias alquilas dos ésteres oleato, linoleato e dos ésteres saturados. O sinal C ( $\delta$ 1,28-1,30)



apresentou correlação adicional ao grupo metileno  $\beta$  carbonila de cadeias alquilas do sinal D ( $\delta$  1,61-1,62), sugerindo assim o desdobramento de um multipletto devido às múltiplas correlações, os sinais A e B são desdobrados em tripletos quando interagem com o grupo metileno do sinal C. Como o sinal D ( $\delta$  1,61-1,62) também se correlaciona com o grupo metileno  $\alpha$  carbonila de cadeias alquilas do sinal F ( $\delta$  2,31-2,33) a ressonância desses núcleos gera um multipletto, avaliando a interação do sinal F com o D, o seu desdobramento gera um tripnetto. Observou-se também que o sinal G ( $\delta$  2,77-2,79) presente nos ésteres linoleato e linolenato, correlacionam-se com os grupos metínicos olefínicos do sinal H ( $\delta$  5,34-5,37) e, que estes também se correlacionam com os hidrogênios E ( $\delta$  1,95-2,13) dos ésteres oleato, linoleato e linolenato. Os sinais de hidrogênio alfa ( $H_\alpha$ ) da unidade do glicerol acoplam com o hidrogênio beta ( $H_\beta$ ), gerando um desdobramento de dupletos de dupletos (dd) para os hidrogênios nas posições sn-1 e sn-3 da unidade do glicerol. Já o  $H_\beta$  quando interage com os  $H_\alpha$  se desdobram em um multipletto.

Figura 12: Ampliação do mapa de contornos COSY da região do grupo metila do éster linolenato

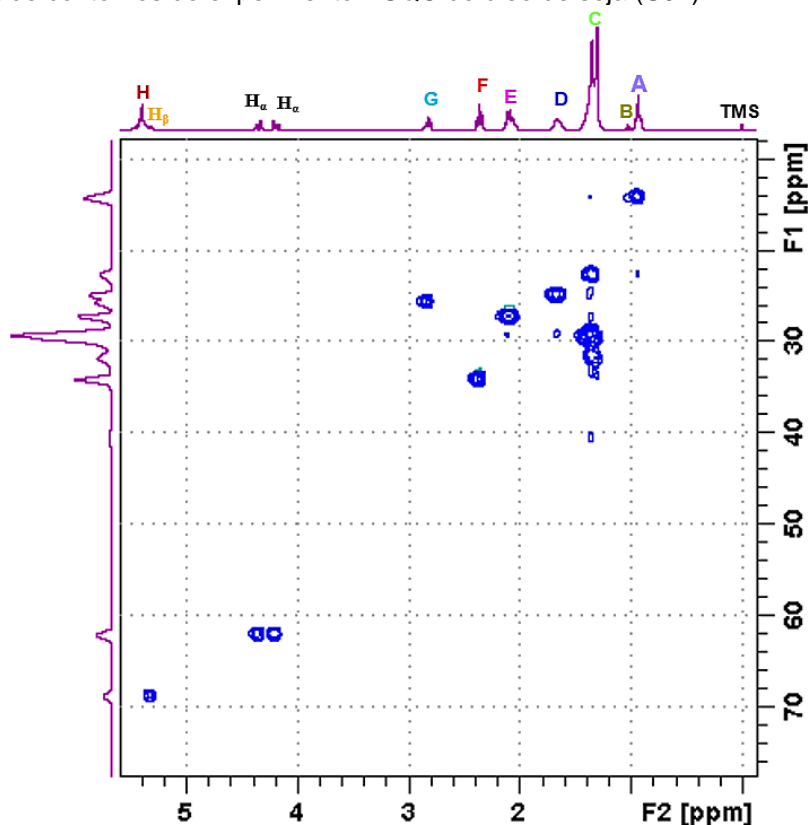


### 5.1.3 Atribuição dos sinais nos espectros de RMN bidimensional heteronuclear (HSQC e HMBC) dos óleos de soja e andiroba.

A atribuição dos deslocamentos químicos dos carbonos dos TAGs foi realizada a partir dos mapas de contorno dos experimentos HSQC (figura 13) e HMBC (figura 14), que também avaliaram os acoplamentos heteronucleares de curta ( $^1J_{CH}$ ) e longa distância ( $^{2,3}J_{CH}$ ), respectivamente.

Os hidrogênios olefínicos em 5,34-5,37 ppm apresentaram correlações  $^1J_{CH}$  com os carbonos 130,06 ppm e 128,04 ppm. O sinal referente ao hidrogênio da posição sn-2 da molécula do glicerol teve correlação  $^1J_{CH}$  em 68,84 ppm, assim como os hidrogênios sn-1 e sn-3 tiveram correlação  $^1J_{CH}$  em 62,12 ppm, caracterizando a presença dos TAGs. Os hidrogênios metilênicos que se encontram entre os grupos olefínicos dos ésteres linoleato e linolenato 2,77-2,79 ppm apresentou correlação  $^1J_{CH}$  em 25,69 ppm e os hidrogênios insaturados encontrados em todos os ésteres com deslocamento em 2,03 ppm possuem correlação  $^1J_{CH}$  em 27,21 ppm. Os hidrogênios  $\alpha$  carbonila em 2,31-2,33 ppm tem correlação  $^1J_{CH}$  em 34,08 ppm e os hidrogênios metilênicos alifáticos mostrou correlações  $^1J_{CH}$  em 29,25 e 31,55 ppm. Os sinais de metilas terminais em ~0,90 e 1,02 ppm apresentaram correlação  $^1J_{CH}$  com 14,13 ppm.

Figura 13: Mapa de contornos do experimento HSQC do óleo de soja (S02)



Os hidrogênios olefínicos em 5,34-5,37 ppm apresentaram correlações  $^nJ_{CH}$  com os carbonos 25,69 ppm, 27, 61 ppm e 29,13-29,52 ppm. O sinal referente ao hidrogênio da posição sn-2 da molécula do glicerol teve correlação  $^nJ_{CH}$  em 62,12 ppm, 172,44 ppm. Os hidrogênios sn-1 e sn-3 tiveram correlação  $^nJ_{CH}$  em 62,12 ppm, 68,90 ppm e 173,25 ppm, caracterizando a presença dos TAGs.

Os hidrogênios metilênicos que se encontram entre os grupos olefínicos dos ésteres linoleato e linolenato 2,77-2,79 ppm apresentaram correlação  $^nJ_{CH}$  em 27,21 ppm, 29,13-29,52 ppm, 127,92 ppm e 130,06 ppm. Os hidrogênios insaturados encontrados em todos os ésteres com deslocamento em 2,03 ppm possuem correlação  $^nJ_{CH}$  em 29,13-29,52 ppm, 31,55 ppm, 127,92 ppm e 130,06 ppm. Os hidrogênios  $\alpha$  carbonila em 2,31-2,33 ppm tem correlação  $^nJ_{CH}$  em 24,83 ppm, 29,13-29,52 ppm, 172,44 ppm e 173,25 ppm; os hidrogênios metilênicos alifáticos mostrou correlações  $^nJ_{CH}$  em 14,13 ppm, 24,84 ppm, 27,21 ppm, 29,13-29,52 ppm e 33,92 ppm. Os sinais de metilas terminais do grupo linoleil apresentaram correlação  $^nJ_{CH}$  com 22,73 e 31,55 ppm, já as metilas terminais do linolenil apresentaram correlação  $^nJ_{CH}$  com 20,56 ppm e 131,8 ppm.

Figura 14: Mapa de contornos do experimento HMBC do óleo de soja (S02)

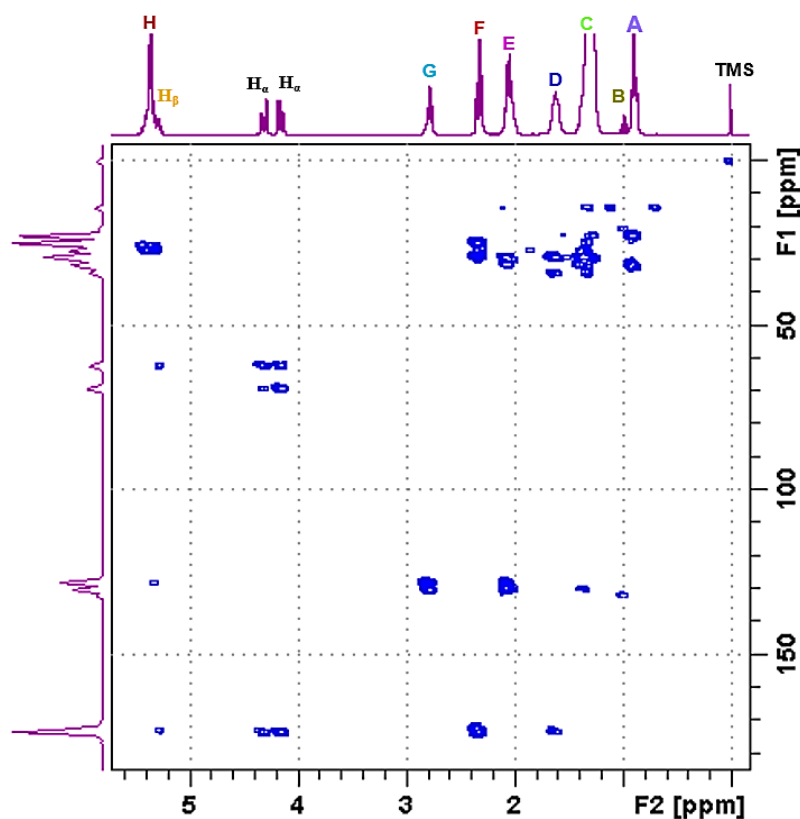


Tabela 5: Atribuição dos deslocamentos químicos dos carbonos de triacilgliceróis dos óleos vegetais por meio dos experimentos de HSQC e HMBC

Carbonos	Sinal	<sup>1</sup> H	Deslocamento químico, δ (ppm)	
			HSQC	HMBC
CH <sub>3</sub> -(CH <sub>2</sub> ) <sub>n</sub>	A	0,86	14,13	22,73 e 31,55
CH <sub>3</sub> -CH <sub>2</sub> -CH=CH-	B	0,92	20,56	20,56 ppm e 131,8 <sup>a</sup>
(CH <sub>2</sub> ) <sub>n</sub>	C	1,28	22,71; 29,25 e 31,55	14,13; 24,84; 27,21; 29,13-29,52 e 33,92
-CH <sub>2</sub> -CH <sub>2</sub> -COOR	D	1,61	24,83 e 131,8 <sup>a</sup>	29,13-29,52; 34,08; 172, 44 <sup>b</sup> e 173,25 <sup>c</sup>
-CH <sub>2</sub> -CH=CH-	E	2,03	27,21	29,13-29,52; 31,55; 127,92 e 130,06
-CH <sub>2</sub> -COOR	F	2,31	34,08	24,83; 29,13-29,52; 172, 44 e 173,25
-CH=CH-CH <sub>2</sub> -CH=CH-	G	2,77	25,69	27,21; 29,13-29,52; 127,92 e 130,06
-CH <sub>2</sub> -OCOR	Alfa	4,22	62,12	62,12; 68,90 e 173,25
>CH-OCOR	Beta	5,28	68,84	62,12 e 172,44
-CH=CH-	H	5,34	130,06 e 128,04	25,69; 27, 61; 29,13-29,52

<sup>a</sup>Linolenil; <sup>b</sup>Carbono da carbonila do éster em sn-2; <sup>c</sup>Carbono da carbonila do éster em sn-1 e sn-3.

## 5.2 Análise multivariada dos dados

O filtro de transformada de Fourier (FT) foi utilizado para colocar os espectros no domínio da frequência, o ajuste da fase e a correção da linha de base foram realizados para minimizar as oscilações de temperatura do equipamento e realizados de forma automatizada, as áreas dos sinais foram obtidas através da construção dos *buckets*. Outro parâmetro que também não foi aplicado, também conhecido como alisamento, foi a multiplicação exponencial ( $I_b = 0$  Hz), pois os espectros deveriam manter suas características originais para não influenciar nos resultados quimiométricos. Todos os espectros foram calibrados pelo deslocamento do sinal do TMS em 0 ppm.

Após o processamento dos espectros no programa TopSpin, realizou-se uma avaliação prévia dos espectros de RMN de <sup>1</sup>H 1D para verificar se existia alguma diferença entre os espectros. Pois, às vezes algumas classificações e discriminações se dão pela má aquisição dos espectros, principalmente as medidas de *shimming*, correção de fase e de base, alisamento, dentre outros. Contudo, não foi identificada nenhuma alteração visual com os parâmetros de aquisição e nem de processamento dos espectros de hidrogênio adotados. Posteriormente, os espectros foram alinhados

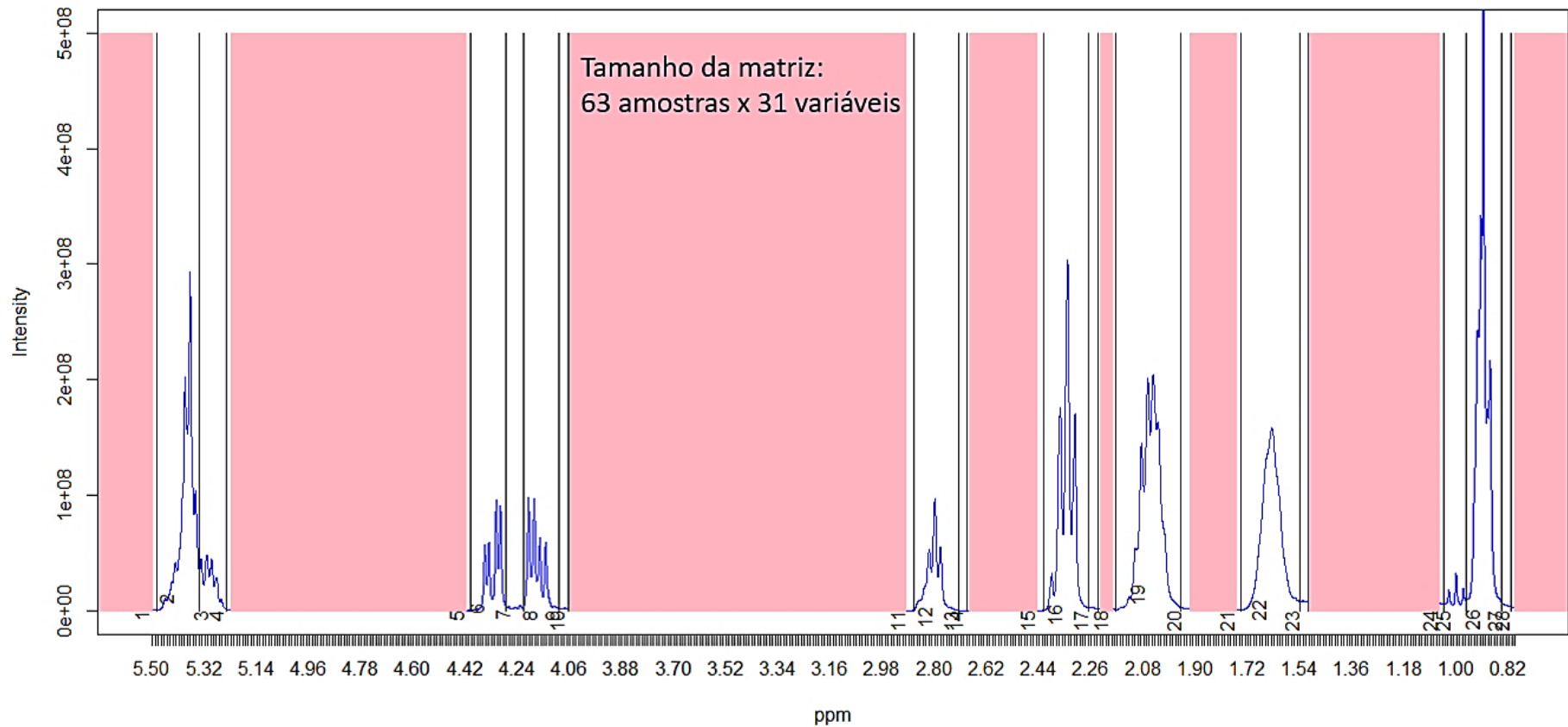
no programa MATLAB através do algoritmo ICOSHIFT visando ajustar as variações de deslocamentos dos sinais do espectro, uma vez que tais flutuações podem estar ligadas às alterações do potencial hidrogeniônico (pH), temperatura e interações intermoleculares (SAVORANI *et al*, 2010).

A transformação das amostras foi realizada através da normalização de todos os sinais do espectro pela área do TMS (0,5% v.v), pois apesar da programação estar programada em 300 K, ainda havia variação da temperatura do espectrômetro. Como o TMS se encontrava com concentração constante, já que se mantinha em um tubo hermeticamente fechado, a normalização das amostras procurou corrigir as variações de temperatura e eventuais erros de concentração das amostras.

No pré-processamento das variáveis dos dados espectrais realizado no RStudio, que já estavam em formato de matriz ( $X_{ij}$ ), o melhor escalamento adotado foi o Pareto, pois os sinais mais intensos poderiam atrapalhar nos resultados das análises quimiométricas e eles são menos suscetíveis aos ruídos. Esse escalamento reduz o peso de sinais intensos fazendo com que a estrutura dos dados se mantenha quase que inalterada, permitindo que todas as variáveis fossem avaliadas com o mesmo peso. Sem esse pré-tratamento, os sinais mais intensos poderiam pesar na discriminação das amostras e levar uma classificação duvidosa.

Antes de realizar os testes de análise exploratória, o tamanho original da matriz era de 63 amostras por 29.012 variáveis, no intervalo de -5,15 a 15,20 ppm. Posteriormente, realizou-se a seleção das variáveis, deixando apenas os sinais de hidrogênio que realmente tinham alguma informação química referente às amostras. Logo, optou-se por trabalhar com uma faixa menor que ia de -1,00 a 8,00 ppm, cujo tamanho da matriz foi reduzido à 63 amostras x 11.413 variáveis; isso sem as regiões de ruído e o sinal do metileno das cadeias alquilas. O metileno foi excluído, porque apesar do pré-tratamento via Pareto, ainda pesava no comportamento das amostras por ser 12 vezes maior que o segundo sinal mais intenso do espectro. Também foi realizada a integração das áreas dos sinais pela criação dos *buckets* ou *binning* automáticos (64 *buckets* ) e otimizados (31 *buckets* ), gerando uma quantidade de variáveis menores. Este tratamento foi utilizado para todos os métodos quimiométricos abordados neste trabalho, tanto supervisionado quanto não supervisionado.

Figura 15: Espectro médio de RMN de  $^1\text{H}$  das amostras do banco 1 (andiroba padrão x andiroba comercial x soja) com a representação dos *buckets* otimizados, matriz 63x31



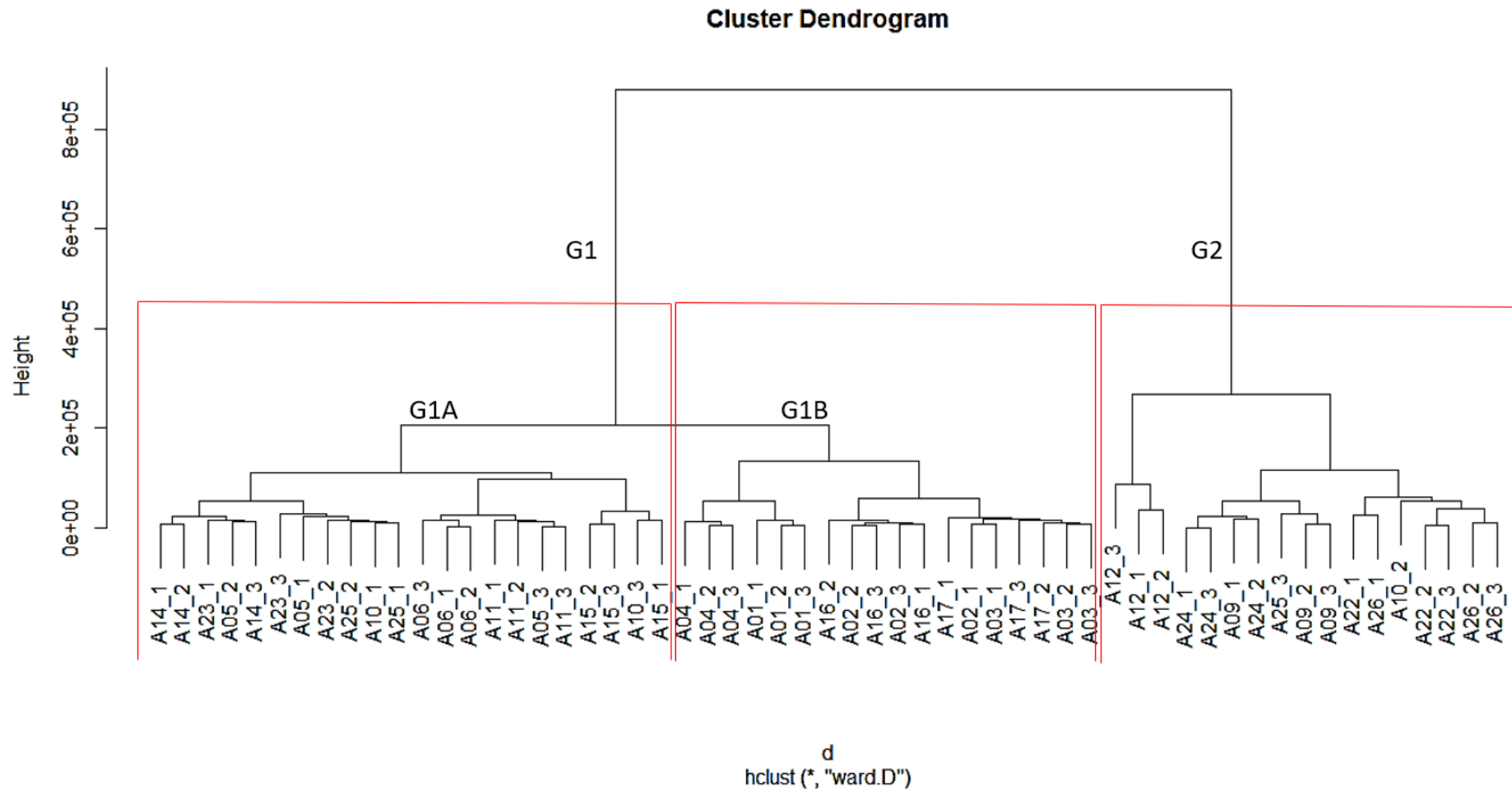
## 5.2.1 Análise exploratória ou não supervisionada

### 5.2.1.1 Análise de agrupamento hierárquico (HCA) do Banco de dados 1 (Andiroba padrão, andiroba comercial e soja)

A análise de agrupamento hierárquico (HCA) interliga as amostras por similaridade, quanto menor a distância entre as amostras ou a altura das linhas que ligam os grupos, maior a semelhança entre elas. A representação gráfica da HCA é feita através de um dendrograma, denominado por diagrama de árvores (figura 16). A seguir, encontra-se a discussão do primeiro dendrograma que avalia o comportamento das amostras de andiroba comercial frente às amostras padrão, sem a presença do adulterante, a soja.

No dendrograma da figura 16, a linha vermelha na horizontal, indica o corte em que foi observado a dissimilaridade dos grupos G1 e G2, apesar da distância dos pontos ligantes não apresentarem uma distância tão proeminente. No grupo G1, pôde-se observar uma subdivisão intragrupo, G1A e G1B, porém a distância dos pontos ligantes intragrupo é mínima, podendo levar a erros de interpretação. O grupo G1B foi formado apenas por amostras de andiroba padrão (A01, A02, A03, A04, A16 e A17) da mesma cidade, obtidos da região de Manaquiri-AM; já o grupo G1A foi formado pelas amostras comerciais que possivelmente são consideradas não adulteradas ou íntegras (A05, A10, A11, A14, A15 e A23), pois as mesmas se agruparam com uma amostra padrão de andiroba (A06) que também foi adquirida na cidade de Manaus-AM, bem como todas amostras comerciais.

Figura 16: Dendrograma da HCA das amostras de andiroba padrão e andiroba comercial



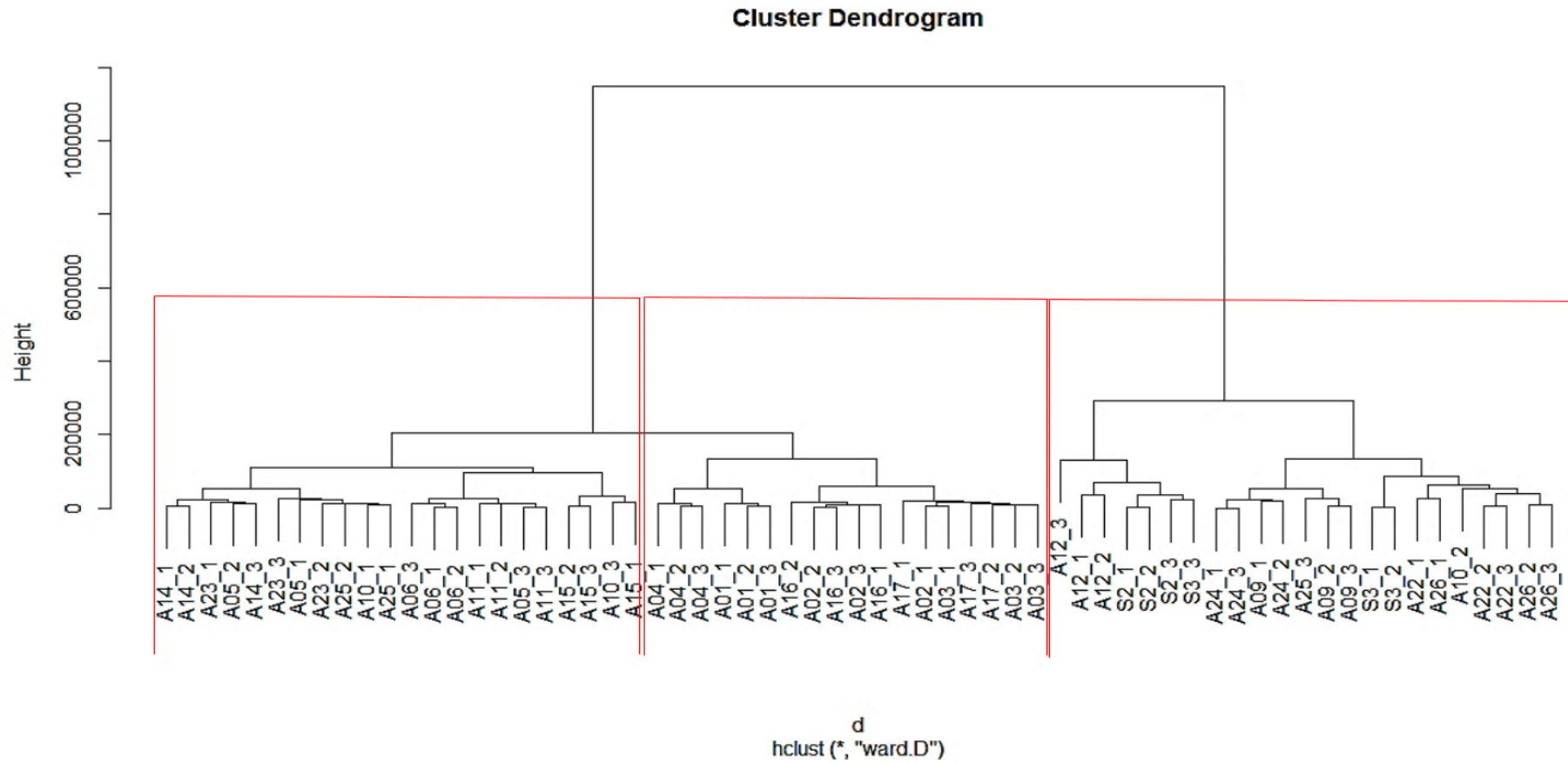


O grupo G2 foi formado apenas por amostras comerciais (A09, A10, A12, A22, A24, A25 e A26), tal comportamento pode ser associado a possíveis adulterações desses óleos, pois os mesmos se distanciaram do grupo G1. O comportamento das réplicas A10\_2 e A25\_3 mostraram possíveis erros experimentais ou as amostras podem ter níveis de adulteração que na análise de agrupamento hierárquico não pôde ser detectado, já que a HCA é uma técnica rígida que define apenas se uma amostra ou réplica pertence ou não pertence ao grupo, ou seja, a HCA avalia as tendências e agrupamentos prévios das amostras por similaridade, distinguindo se a amostra pertence ou não pertence ao grupo.

Como as adulterações em óleos vegetais geralmente são feitas com óleo mais baratos e de fácil acesso, optou-se por realizar a mesma análise de agrupamento com as amostras da figura 16 mais o óleo de soja, pois o mesmo apresenta baixo custo e composição semelhante a cadeia graxa do óleo de andiroba, variando apenas as proporções de ácidos graxos livres e triacilgliceróis (exceto o ácido linolênico).

Na segunda HCA (figura 17) foram observados os mesmos comportamentos de similaridade intragrupo e de dissimilaridade entre os grupos G1 e G2, mesmo com a inclusão das amostras de soja no banco de dados. Outra observação que se mostrou muito pertinente a esta análise exploratória foi o comportamento das amostras do grupo G2 de se agruparem juntamente com as amostras de soja. Logo, há uma probabilidade muito grande das amostras comerciais de andiroba do grupo G2 terem sido adulteradas com óleo de soja, pois segundo o dendrograma dessas amostras possuem características mais próximas do óleo de soja do que do óleo de andiroba.

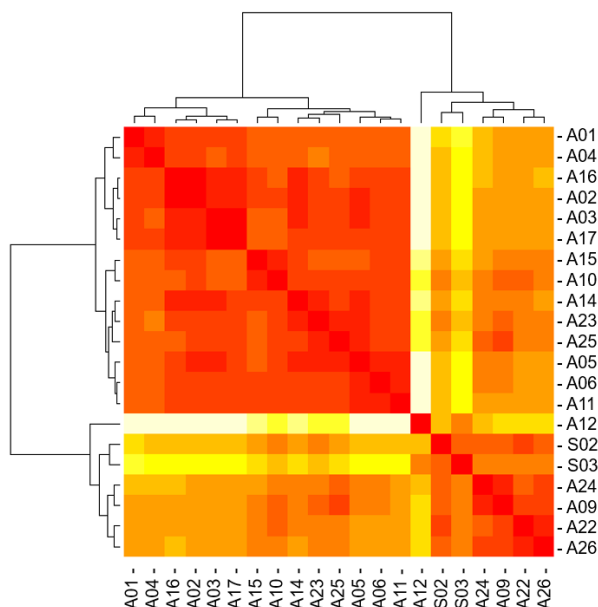
Figura 17: Dendrograma da HCA das amostras andiroba padrão e andiroba comercial versus soja comercial



### 5.2.1.2 Análise de agrupamento Fuzzy C-Means (FCM) do Banco de dados 1 (andiroba padrão, andiroba comercial e soja)

Como a HCA avalia se uma amostra pertence ou não pertence a um grupo, decidiu-se realizar a análise exploratória FCM, que também é hierárquica, porém indica ainda o grau de similaridade de uma amostra com um ou mais grupos. Antes de realizar o FCM foi necessário avaliar o mapa de calor (*heatmap*) da HCA realizada anteriormente (Figura 17), pois ele indica se existe alguma dissimilaridade intragrupo através das cores formadas no mapa. A partir daí, pôde-se estimar o número de agrupamentos (cluster), cujo parâmetro é representado pela letra K. O mapa de calor compara o espelhamento do dendrograma das amostras da HCA.

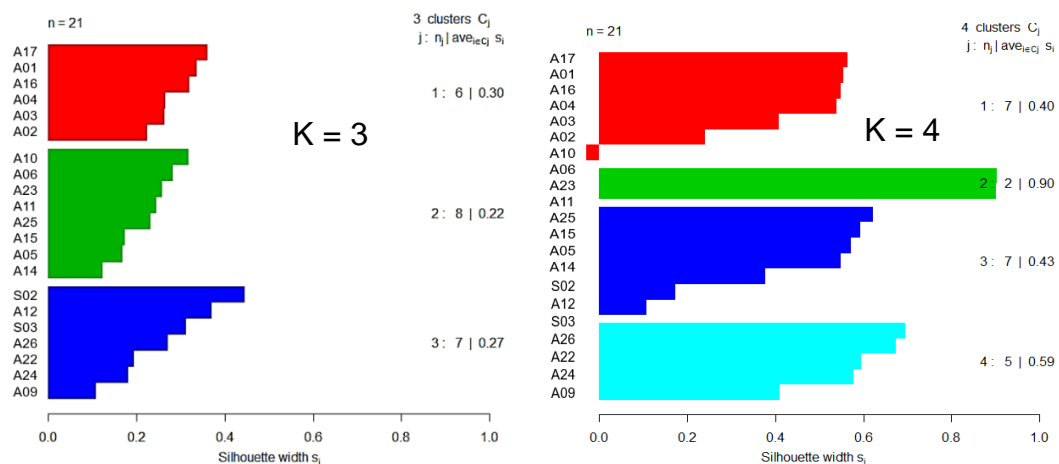
Figura 18: Mapa de calor da média das réplicas das amostras de andiroba padrão e comercial *versus* o óleo de soja comercial



O mapa de calor (figura 18) mostrou que, além dos dois grupos formados G1 (Cor mais intensa) e G2 (cor mais clara), aparecem duas cores predominantes em cada grupo. Em que G1, tem cores vermelho e laranja escuro podendo ser representado como G1A, também tem uma cor laranja um pouco mais clara que pode ser representado como G1B e em G2 as cores amarelo escuro e claro. Também foi observado um comportamento anômalo da amostra A12 (cor branca). Assim, o parâmetro de número de grupos (K) que fora avaliado ficou entre 2 a 4. Após os testes com a mudança dos valores de K, observou-se que a formação de 3 grupos era o mais condizente (conforme o gráfico em silhueta representado na figura 19, pois o coeficiente de *Dunn* ficou em torno de 0,9120 (valor máximo assumido é 1) quando a

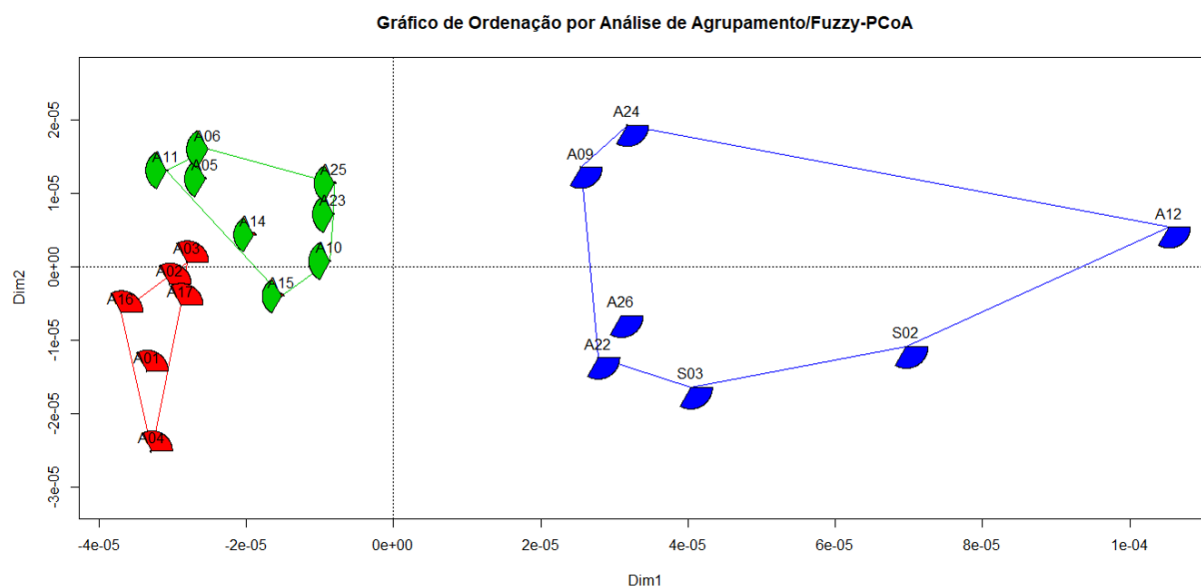
constante de fuzzificação (*membership*) estava em 1,25; em que, abaixo de 1 pode ser considerado apenas uma HCA, acima de 1 avalia o grau de pertencimento das amostras em outros grupos, assim a constante de *Dunn* deve estar entre 0,8-1,0).

Figura 19: Representação gráfica em silhueta da avaliação do melhor agrupamento encontrado  $K = 3$



O segundo gráfico (figura 20) gerado na análise de FCM é o gráfico de ordenação (PCoA), ele é semelhante a análise de componentes principais (PCA), contudo a PCoA não indica quais variáveis influenciam no agrupamento das amostras, também não mostra a relação entre as variáveis e as componentes principais, avalia apenas a similaridade entre as amostras.

Figura 20: Representação gráfica da análise de *Fuzzy C-Means* da média das réplicas das amostras do banco 1



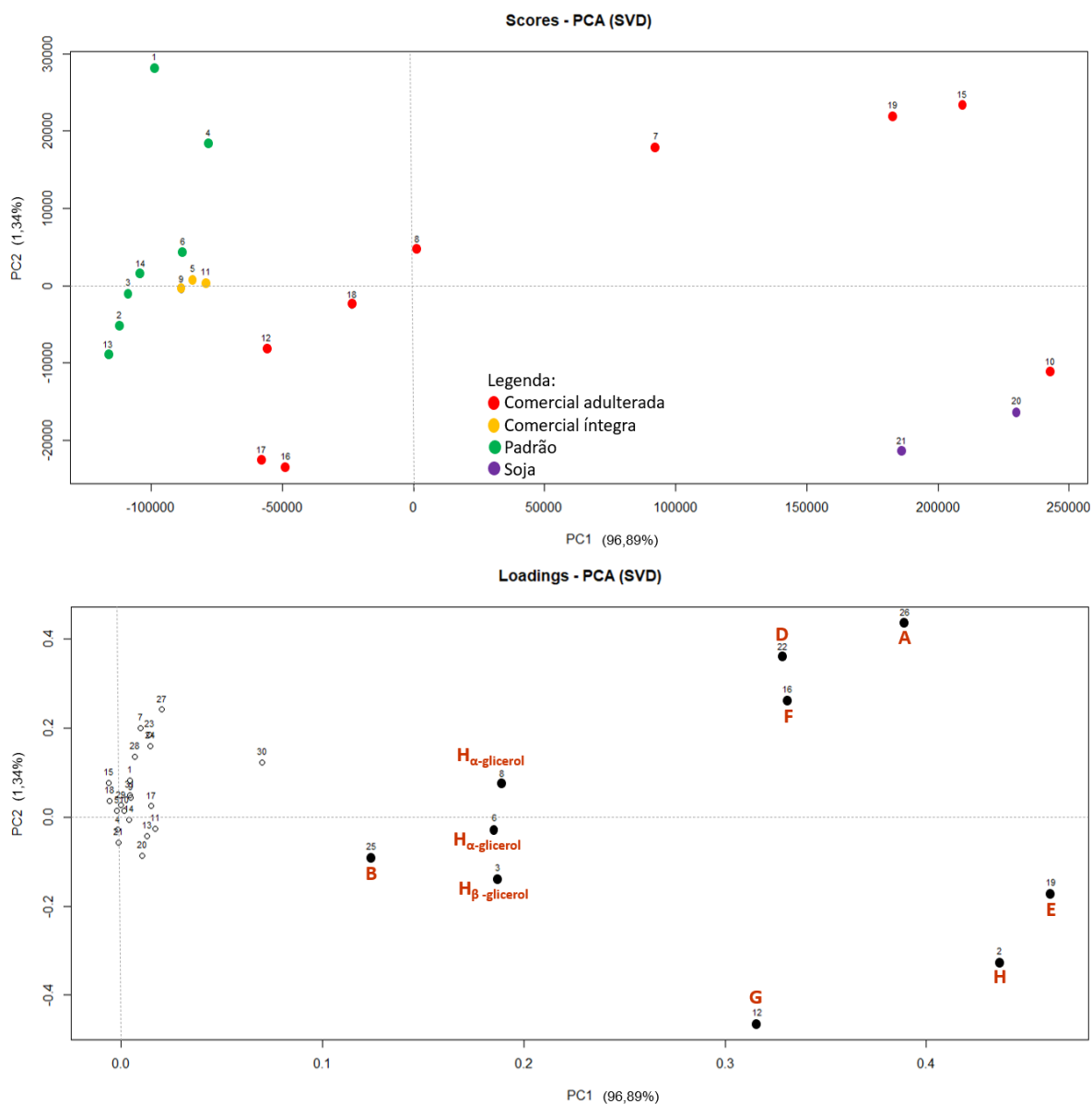
Na gráfico de PCoA (figura 20), pôde-se observar a subdivisão intragrupo do grupo G1, denominada anteriormente pelo mapa de calor como G1A e G1B. No leque de cor vermelha (subgrupo G1B) o grupo foi formado apenas por amostras de andiroba padrão (A01, A02, A03, A04, A16 e A17), todas pertencentes à mesma cidade de origem, Manaquiri-AM. No leque de cor verde (subgrupo G1A) o agrupamento se deu com algumas amostras comerciais de andiroba (A05, A10, A11, A14, A15 e A23) e uma amostra padrão (A06), todas obtidas na cidade de Manaus-AM. À priori, as amostras de cor verde em PCoA foram consideradas não adulteradas ou íntegras, já que se agruparam com uma amostra padrão de andiroba (A06) e ficam próximas das demais amostras padrão, que se encontram na cor vermelha do gráfico de ordenação. Apesar das amostras terem sido separadas pela região de onde foram produzidos ou vendidos, ainda não se pode concluir o agrupamento sem a avaliação das variáveis que pesaram para esse agrupamento. Além do que, existem divergências de informações dos vendedores quanto à origem dos óleos comerciais que são revendidos nas feiras da cidade de Manaus. Ainda no gráfico de PCoA, verificou-se que o leque de cor azul (grupo G2 dos dendrogramas) foi formado por algumas amostras de andiroba comercial (A09, A12, A22, A24 e A26) e as amostras de soja (S02 e S03). Isso indica suspeitas de adulteração dessas amostras com óleo de soja.

### *5.2.1.3 Análise exploratória por PCA robusta do Banco de dados 1 (andiroba padrão, andiroba comercial e soja)*

Após o pré-tratamento dos dados foram realizados testes preliminares com a PCA comum, utilizando o algoritmo SVD, contudo como este método é sensível a *outliers*, sempre que era utilizado o gráfico da distância ortogonal e a distância de *scores* para avaliar se essas amostras eram realmente amostras anômalas, assim que eram retiradas as que eram detectadas como *outliers*, surgiam novas amostras com comportamento anômalo e continuava assim até não ter mais amostras para serem retiradas. Analisando a figura 21, em que os gráficos de *scores* e *Loadings* estão representados, verificou-se que as amostras no gráfico de *scores* não podem ser justificadas através do gráfico de *Loadings*, pois não conseguem discriminar as variáveis para estas amostras em PC1 negativo, bem como as amostras dispersas

em PC1 positivo. Então, optou-se por realizar a PCA robusta pelo método “mad” (desvio médio absoluto), pois não é sensível a *outliers* uma vez que ela se move na direção da maior quantidade de dados pelo desvio médio e não apenas à variância.

Figura 21: Gráficos de *scores* e *Loadings* PC1 X PC2 da PCA (SVD)

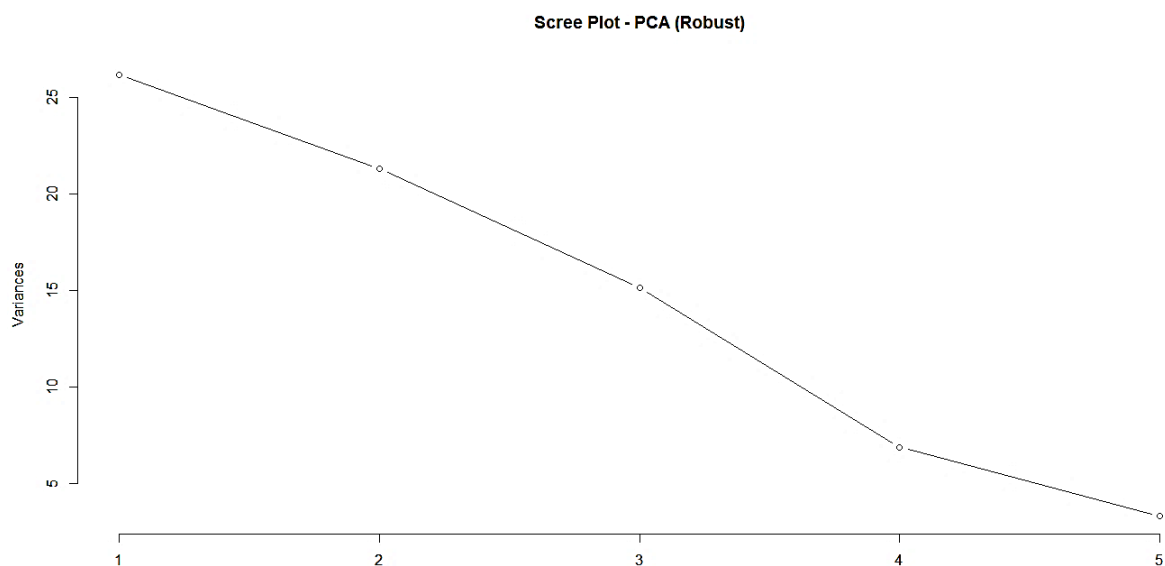


Logo, optou-se por realizar a PCA robusta (figura 23), que assim como a PCA comum, além de mostrar as tendências ou comportamento das amostras, também discrimina as possíveis variáveis responsáveis pela disposição das amostras de maneira exploratória. O diferencial da PCA robusta é que ela utiliza a mediana nos cálculos dos autovetores e considera todas as amostras de um conjunto, ou seja, ela não é sensível a comportamentos anômalos. Para essa análise foi utilizada a mesma

matriz da HCA, com o intuito de realizar uma comparação dos resultados obtidos no quesito 5.2.1 e 5.2.2.

O gráfico de *Scree Plot* ou gráfico de desempenho (figura 22) auxiliou na escolha de quantas componentes foram necessárias para representar a maior variância explicada da análise de componentes principais. Assim, as componentes principais são ordenadas de forma decrescente de representatividade, da mais significativa para a menos significativa. Neste caso foram utilizadas 3 componentes: PC1(37,63%), PC2 (30,64%) e PC3 (21,81%); totalizando 90,08% da variância explicada. Outro critério de Pearson também foi utilizado, nele as componentes principais devem compreender pelo menos 80% da informação total da descrição dos dados (variabilidade total). Por essa razão, as componentes principais PC4 e PC5 não foram utilizadas.

Figura 22: Gráfico de desempenho das variâncias das componentes principais (PC1 a PC5) e suas contribuições para a PCA robusta



As amostras de óleo de andiroba foram divididas em três grupos: padrão, comercial íntegra e comercial adulterada; os códigos individuais e suas classificações podem ser acompanhados através do quadro 6. Outro grupo também foi utilizado para esse banco de dados, adicionaram-se duas amostras de óleo de soja comercial (S02 e S03), cujo objetivo foi para o controle de adulteração das amostras comerciais. Os critérios considerados na escolha do óleo de soja foram devidos ao custo e acessibilidade, além de possuir uma composição química não muito diferente à

andiroba, variando apenas as proporções dos TAG, ésteres e AG, conforme discutido anteriormente na HCA.

A Análise de Componentes Principais (PCA) robusta além de ter reduzido a dimensionalidade dos dados, ajudou a avaliar o comportamento do banco de dados 1 mediante às suas variáveis. Portanto, a PCA permitiu discriminar os grupos de amostras de óleos através das variâncias de cada componente principal (PC), que está diretamente relacionada ao perfil da cadeia graxa desses óleos. Nesta análise de PCA foram utilizadas 60 amostras (linhas) e 34 variáveis (colunas). A redução de variáveis se deu pela seleção de sinais de hidrogênio que apresentam informações químicas que estão associadas a composição da cadeia graxa dos óleos.

Quadro 6: Identificação das amostras de óleo de andiroba após a análise exploratória por PCA robusta do Banco de dados 1

<b>Código</b>	<b>Identificação das amostras de andiroba</b>
A01	Padrão
A02	Padrão
A03	Padrão
A04	Padrão
A06	Padrão
A16	Padrão
A17	Padrão
A05	Comercial íntegra
A11	Comercial íntegra
A14	Comercial íntegra
A09	Comercial adulterada
A10	Comercial adulterada
A12	Comercial adulterada
A15	Comercial adulterada
A22	Comercial adulterada
A23	Comercial adulterada
A25	Comercial adulterada
A26	Comercial adulterada

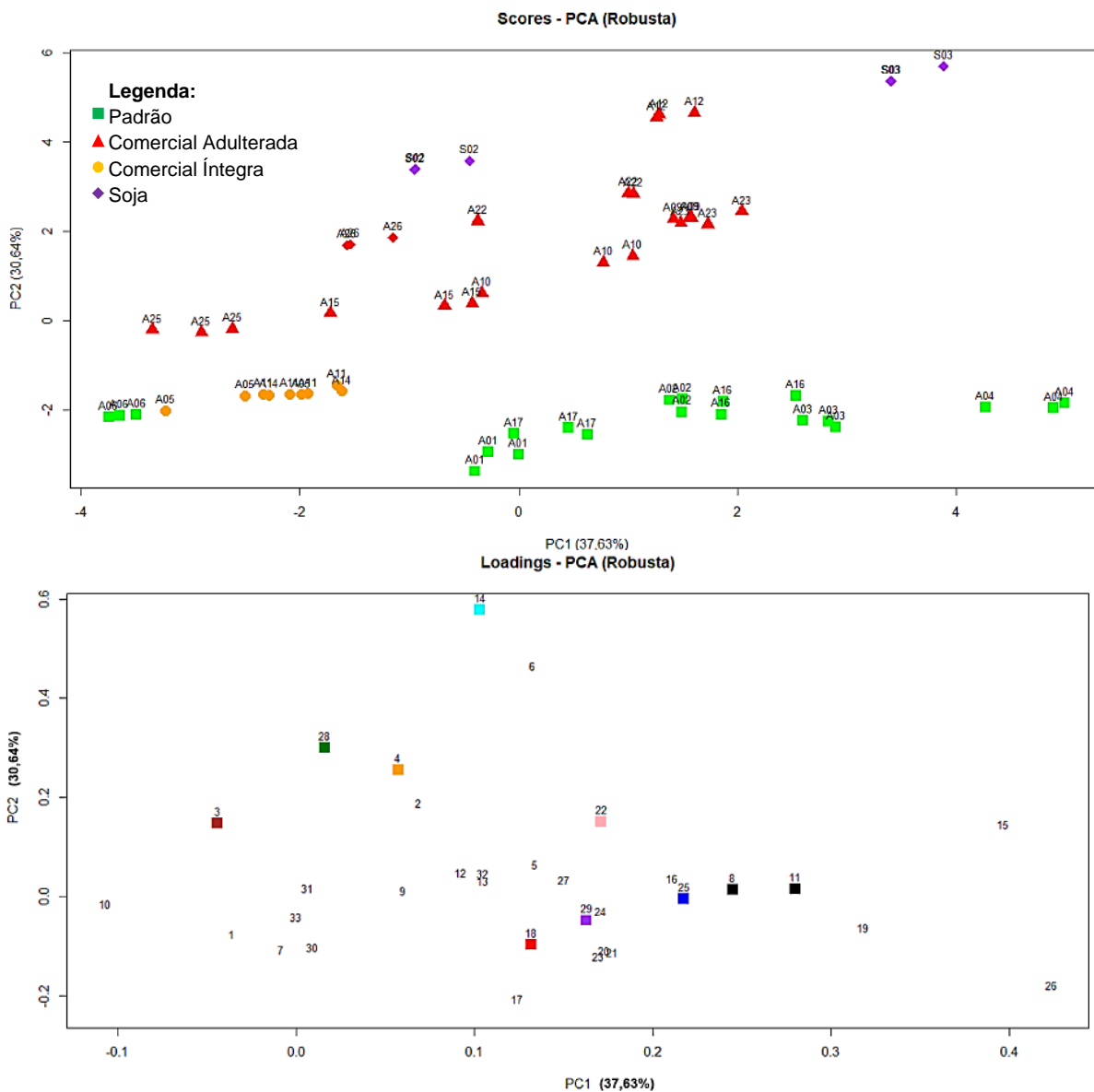
\*Soja 02 e Soja 03

A figura 23 apresenta os gráficos de *scores* e *Loadings* da PCA robusta do **Banco de dados 1** (PC1XPC2). Através do gráfico de *scores* (figura 23) pôde ser avaliado o comportamento das amostras comerciais frente aos padrões de andiroba e soja, através da variância explicada das componentes PC1 (37, 63%) e PC2 (30,64%), cuja variância acumulada explicada foi de 68,27%. Apesar da componente PC1(37,63%) apresentar maior variância explicada, a componente PC2 (30,64%) foi



a responsável pela discriminação dos grupos. Em que, na PC2 negativa foi identificado o grupo de amostras de andiroba padrão (cor verde) e amostras

Figura 23: Gráfico de *scores* e *loadings* PC1 (37,63%) versus PC2 (30,64%) da PCA (Robusta) das amostras de óleo de andiroba



comerciais íntegras (cor laranja); na componente PC2 positiva foi identificado o grupo de amostras comerciais adulteradas (cor vermelho) e soja (cor roxo).

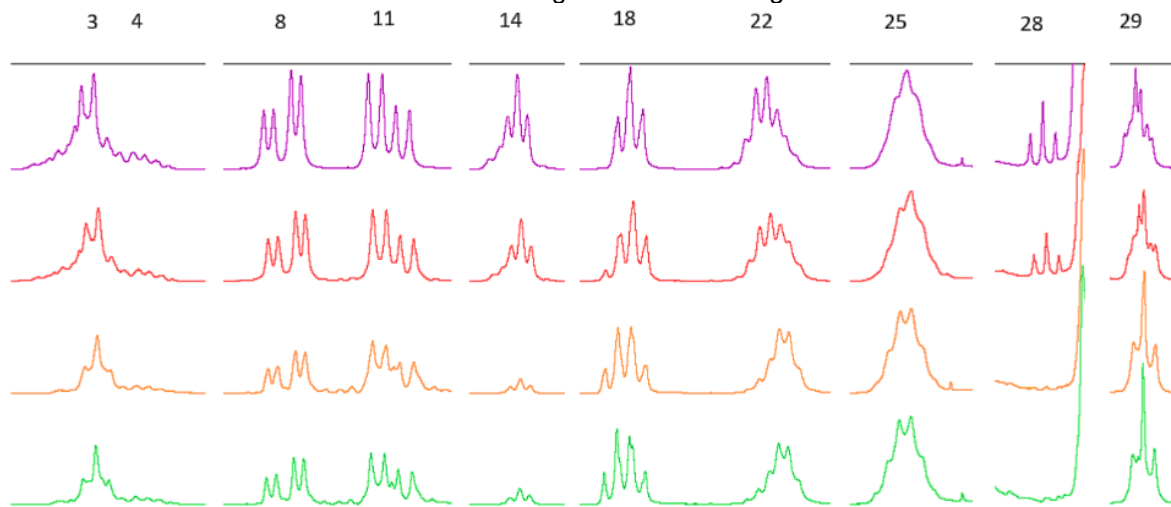
Acompanhando o gráfico de *scores* PC1 versus PC2 foi possível identificar as seguintes tendências: as amostras de andiroba padrão (cor verde) se separam das amostras comerciais e soja, com exceção da amostra padrão A06 (16,17,18) que se agrupa com as amostras comerciais A05, A11 e A14. Logo, pôde-se inferir que as amostras comerciais A05, A11 e A14 se comportam como amostras íntegras (cor laranja), já que ficaram próximas ao padrão de andiroba A06 (cor verde) e distante

das outras amostras comerciais de andiroba (cor vermelha). Outra observação das amostras padrão de andiroba foi que a A06 ficou distante das demais porque foi coletada em Manaus-AM e foi extraída em 2008. As amostras padrão A01, A02, A03, A04, A16 e A17 (cor verde) foram todas adquiridas no município de Manaquiri-AM em 2012, por essa razão todas permaneceram praticamente em PC1 positiva, elas só não ficaram mais próximas porque cada amostra é de um produtor distinto, ou seja, a extração e a localização também pode ser um critério de autenticação do óleo de andiroba, como é feito ao óleo de azeite.

No outro grupo, localizado em PC2 positiva, as amostras comerciais de andiroba foram denominadas de comerciais adulteradas (cor vermelha) porque se encontram no lado oposto às amostras padrão e ficaram próximas da amostra de soja S02 (cor roxo). Contudo, como não se pode afirmar a extensão de adulteração realizada, já que essas amostras também se discriminaram de forma diferente. As amostras A15 e A25 ficam próximas ao eixo zero, porém A15 também se agrupa próxima à soja, considerando-a amostra comercial adulterada. As amostras comerciais adulteradas A09, A10, A12, A22 e A26 são as que se aproximam mais das amostras de soja, o que indica que essas amostras podem ter sido adulteradas grosseiramente e sem nenhum critério de proporção de óleo de andiroba e adulterante.

Para avaliar as variáveis das componentes principais, utilizou-se o gráfico de *Loadings* (figura 23) que associa as variáveis originais que foram responsáveis pela discriminação dos grupos discutidos no gráfico de *scores* (figura 23). As variáveis originais são as áreas dos sinais dos hidrogênios que caracterizam as substâncias responsáveis pela composição da cadeia graxa dos óleos, no gráfico de *loadings* as variáveis são identificadas conforme a numeração da figura 24 da página a seguir.

Figura 24: Sinais de RMN de  $^1\text{H}$  das amostras de óleo de andiroba padrão (verde), andiroba comercial íntegra (laranja), andiroba comercial adulterada (vermelho) e soja (roxo). Os números acima dos sinais são os números dos *buckets* encontrados nos gráficos de *Loadings*



A interpretação dos dados também contou com o auxílio da literatura por meio das informações da cadeia graxa dos óleos de andiroba e soja obtidos por cromatografia gasosa (Tabela 3). Em CG os óleos passam por reações de derivatização para torná-los compostos mais voláteis, por essa razão a tabela 3 faz referência aos ácidos graxos, mas nada impede de comparar aos ésteres presentes nos TAG encontrados nos óleos em questão.

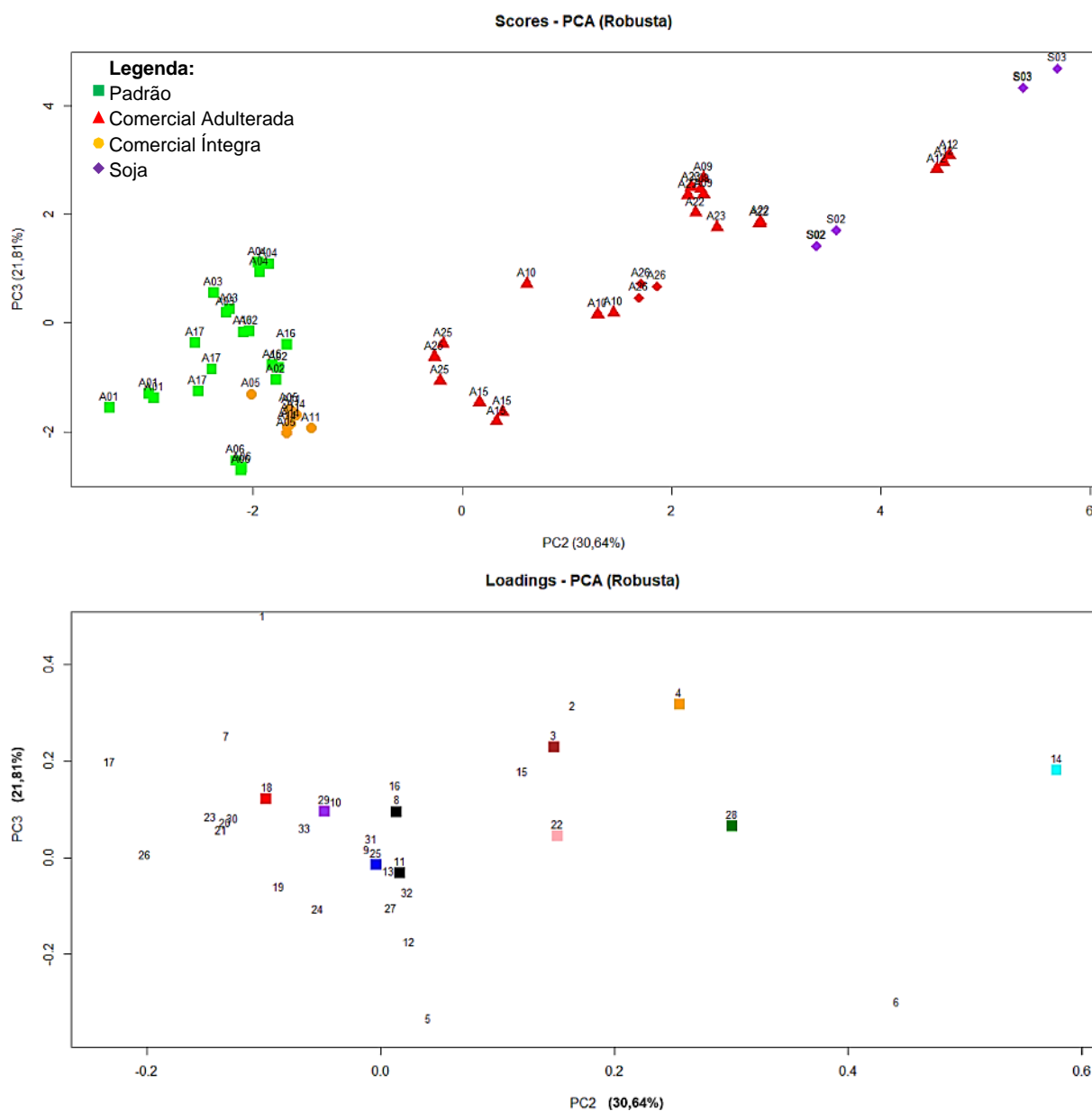
Observando o gráfico de *Loadings* em PC2 negativa, os *buckets* 29, 25, 18, 11 e 8 foram os responsáveis pela discriminação dos grupos de amostras de andiroba padrão e andiroba comercial íntegra. O *bucket* 29 está associado aos AG e ésteres que possuem uma metila de cadeia alquila, principalmente os saturados, uma vez que a andiroba possui 40,2% de cadeia saturada e a soja possui apenas 16%. O *bucket* 18 representa o sinal metileno  $\alpha$  carbonila de cadeias alquilas, nessa região ocorre a sobreposição de dois tripletos do metileno do AG e dos TAG, dando um aspecto de quarteto. Neste caso, a diferença se dá pela característica físico-química dos óleos, uma vez que o óleo de andiroba possui acidez mais elevada que o óleo de soja, por isso que o “quarteto” é mais pronunciado nos espectros de andiroba da amostra A06 (a mais antiga). Os *buckets* 25 e 18 também estão correlacionados às cadeias saturadas da composição química do óleo de andiroba. Portanto, a classificação do grupo de andiroba padrão e comercial-íntegra se dá principalmente pelos ácidos palmítico e esteárico.

Em PC2 positiva, os *buckets* 28, 22, 14, 4 e 3 foram os responsáveis pela discriminação das amostras comerciais adulteradas e a soja. O *bucket* 28 é característico da metila do éster linolenato (ômega 3) que está presente apenas no óleo de soja (8%). O *bucket* 22 é o sinal metileno  $\alpha$  olefínicos de ésteres insaturados, logo como a soja possui 84% de ésteres insaturados e a andiroba apenas 59,81%, a área desse sinal para amostras de soja e que contenham soja são relativamente maior e com a multiplicidade não muito bem resolvida do que a área dos sinais do óleo de andiroba (figura 24). O *bucket* 14 é específico do metileno que fica entre olefinas dos ésteres linoleato e linolenato, que correspondem 59% da cadeia insaturada do óleo de soja e apenas 8,8% de linoleato na andiroba. O *bucket* 3 também é associado aos insaturados, pois ele é o hidrogênio das insaturações, assim como discutido anteriormente, podem ser altamente correlacionados ao óleo de soja.

Analisando a figura 25 do gráfico de *scores* das componentes PC2 (30,64%) e PC3 (21,81%), observou-se uma melhor separação dos grupos discutidos anteriormente, cuja variância explicada acumulada foi de 52,45%. A componente PC2 continua sendo necessária para discriminar as amostras padrão e comerciais íntegras das amostras comerciais adulteradas e soja. As amostras padrão de andiroba (cor verde) e comerciais íntegras (cor laranja) seguem um agrupamento em PC2 negativa, cuja amostra padrão A06 e as amostras comerciais íntegras ficaram mais próximas do que quando avaliadas no gráfico de *scores* PC1 *versus* PC2. Contudo, ainda é possível verificar um afastamento da amostra A06 devido ao local de coleta e tempo de extração.

Em PC2 positiva, encontram-se as amostras comerciais adulteradas e a soja. A amostra comercial de andiroba A12 ficou entre as duas amostras de soja, o que indica alto grau de adulteração. A amostra A09, A22 e A23 ficaram próximas às amostras S02 referente ao óleo de soja, o que poderia indicar média a alta adulteração por óleo de soja. Com exceção da amostra A09, as amostras A22 e A23 foram compradas na feira Adolpho Lisboa, na zona sul da cidade de Manaus. A amostra A10 e A26 se agruparam próximas as amostras consideradas de média a alta adulteração e são de feiras distintas. As amostras A15 e A25 formaram outro grupo de amostras de andiroba comerciais adulteradas em PC3 negativa, ambas foram adquiridas na feira do Zumbi em bancas e dias diferentes, na zona leste da cidade de Manaus, levando a crer que o mesmo fornecedor vendeu o mesmo óleo a esses revendedores.

Figura 25: Gráficos de *scores* e *loadings* PC2 (30,64%) versus PC3 (21,81%) das amostras de óleo de andiroba padrão, andiroba comercial adulterada, andiroba comercial íntegra e soja



No gráfico de *Loadings* (figura 25), os *buckets* 28, 14 e 4 são os que discriminam as amostras mais adulteradas e os *buckets* 3 e 22 discriminam as demais amostras comerciais adulteradas. Como descrito na avaliação de PC1 x PC2, a cadeia graxa insaturada é a maior responsável pela discriminação desse grupo, principalmente pelos ácidos graxos linoleico (C18:2 *cis*-9,12) e linolênico (C18:3 *cis*-6,9,12), exceto o ácido oleico. Para o grupo de amostras padrão e amostras comerciais íntegras os *buckets* responsáveis são 29, 25, 18, 11 e 8 típicos de ácidos graxos saturados que predominam na composição do óleo de andiroba.

### 5.2.2 Análise classificatória ou supervisionada pelo método dos mínimos quadrados parciais (I-PLS) do Banco de dados 2.

Durante as análises espectroscópicas e quimiométricas do **Banco de dados 1**, observou-se o comportamento das amostras de andiroba comercial frente ao óleo de andiroba padrão e a soja. A composição da cadeia graxa desses óleos foi determinante para classificar o óleo de andiroba comercial em íntegra e adulterada, apesar dos sinais serem praticamente os mesmos, com exceção do sinal do ácido linolênico. A amostra A05 fora classificada como íntegra pela análise quimiométrica do **Banco de dados 1**), observando minuciosamente a área dos sinais do espectro de RMN de  $^1\text{H}$  dessa amostra, notou-se que o seu perfil espectroscópico contém pequenos traços de adulteração, quando comparado ao perfil das amostras comerciais adulteradas. Por essa razão foi proposto trabalhar com misturas de óleo de andiroba padrão com óleo de soja, a fim de estimar e classificar a adulteração que foi realizada para cada amostra comercial, desde adulteração fina a adulteração grosseira, pois além de classificar as amostras o I-PLS indica o percentual de adulteração.

Os espectros de RMN de  $^1\text{H}$  do **Banco de dados 2** foram obtidos com os mesmos parâmetros de aquisição e processamento, bem como o tratamento das amostras e pré-tratamento das variáveis adotado ao **Banco de dados 1**. Contudo, vale ressaltar que neste banco de dados foi necessário utilizar a média das réplicas para melhor visualização dos dados, pois como as amostras ficaram muito próximas umas das outras, com a presença de todas as réplicas a visualização ficava bastante dificultosa. Como as amostras A11 e A14 não apresentaram perfil de adulteração, não foram incluídas no modelo do I-PLS, estas duas amostras foram as únicas obtidas de empresas certificadas com CNPJ.

#### 5.2.2.1 Modelo de calibração

Para o desenvolvimento do modelo de calibração foram utilizadas as amostras da tabela 6. Com isso foi possível obter o gráfico de validação cruzada que avalia a raiz quadrada do erro médio quadrático da predição (RMSEP) e o número de variáveis latentes (componentes) necessárias para o modelo de calibração que fora utilizado para prever a adulteração de amostras comerciais de andiroba (compradas em feiras

e comércios). Inicialmente, utilizou-se 5 componentes para o modelo, contudo o gráfico da figura 26 mostrou que com apenas 2 componentes já seriam suficientes para descrever o modelo de calibração, já que em 3 componentes permanece quase que a mesma contribuição. Apesar do menor valor de RMSEP indicar o número de variáveis latentes a serem usadas, em que a componente 2 correspondeu a um valor de 5,653 e a componente 3 correspondeu a 5,356; quando é observada a componente 3 para 4, o valor subiu para 7,066. Portanto, realizou-se o teste com 2 e 3 componentes. Outra observação é que os valores de validação cruzada (CV) e validação cruzada ajustada são praticamente as mesmas, o que demonstra não haver sobreajuste do modelo, porém o valor de RMSEP foi de 33,1%.

Figura 26: Gráfico de avaliação do erro médio quadrático de predição *versus* variáveis latentes (componentes)

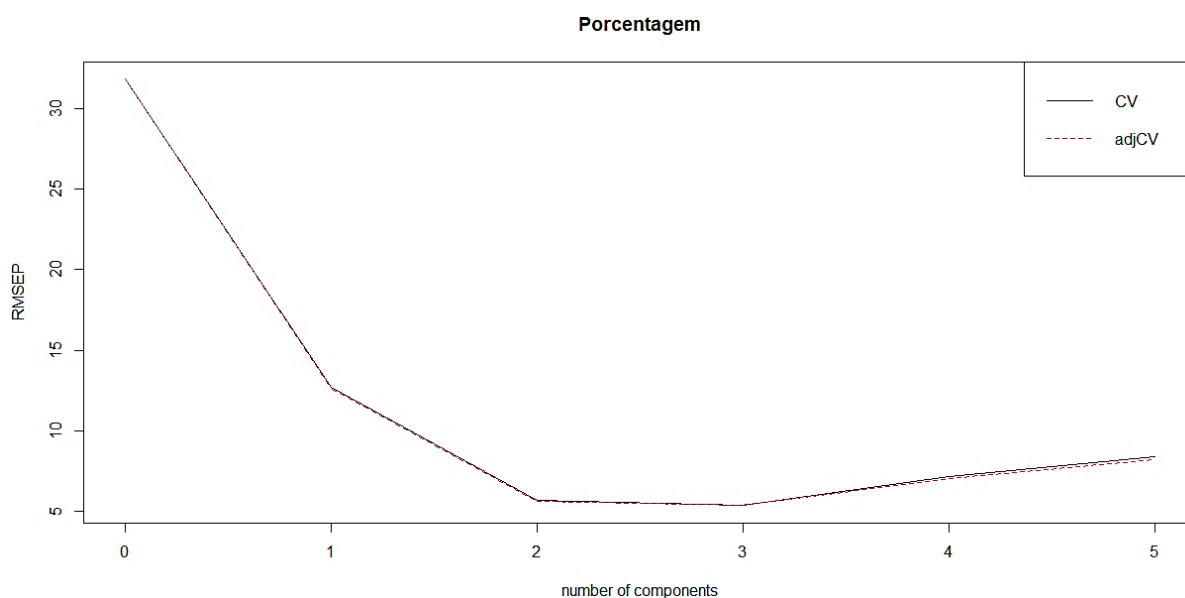
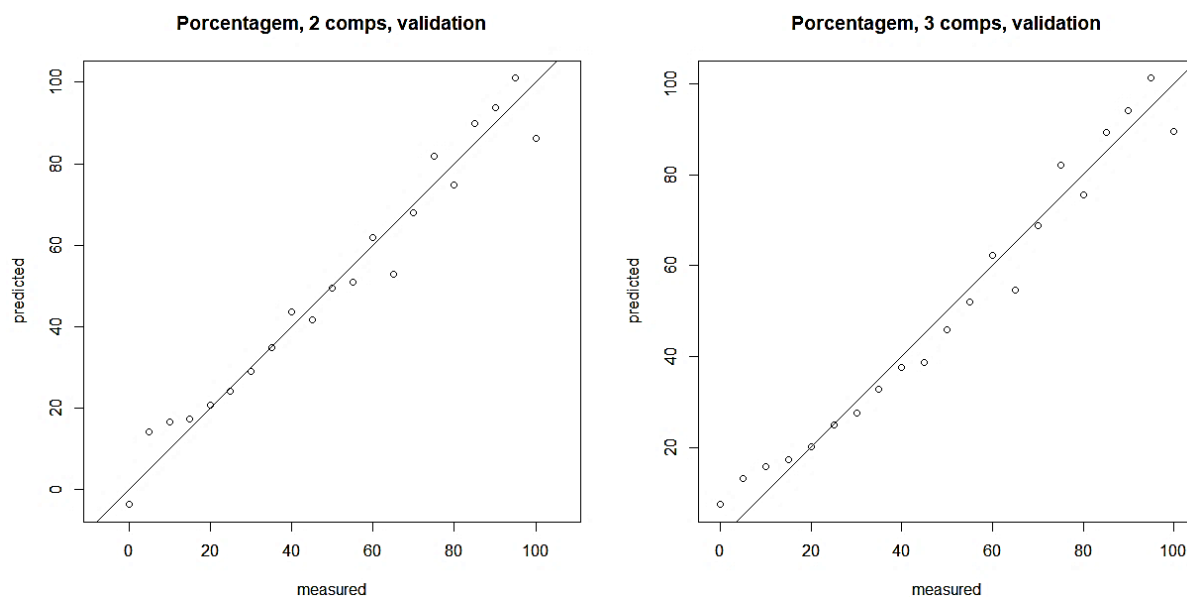
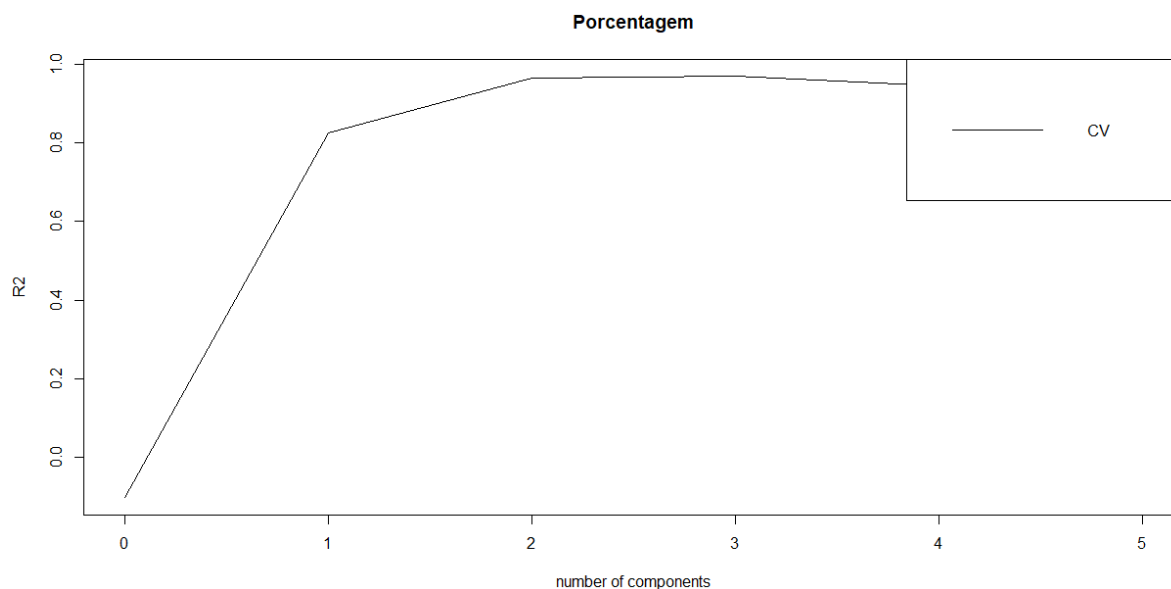


Figura 27: Correlação entre os valores de soja predito (*predict*) pelo modelo e adicionados às misturas (*measured*)



Como foi verificado que apenas duas ou três componentes poderiam ser utilizadas na calibração, por conseguinte foi plotado o gráfico de correlação entre os valores de soja nas misturas e os teores preditos pelo modelo (figura 27).

Figura 28: Gráfico de seleção de modelo



A correlação descrita nos dois gráficos da figura 27 é praticamente a mesma, por essa razão foi decidido plotar o gráfico de seleção do modelo, pois ele compara a modelagem e o poder de predição para determinar o número de componentes a utilizar no modelo. Um modelo excelente de capacidade preditiva normalmente tem uma inclinação em 1 ou muito próximo de 1, o que pode ser visualizado no gráfico de



seleção de modelo (figura 28), cuja inclinação está próxima ao  $R^2 = 0,9683$  e sua inclinação indica que são necessárias apenas duas componentes para o modelo, podendo ser testado com 3, apesar da pouca diferença entre as duas.

### 5.2.2.2 Validação do Modelo de I-PLS

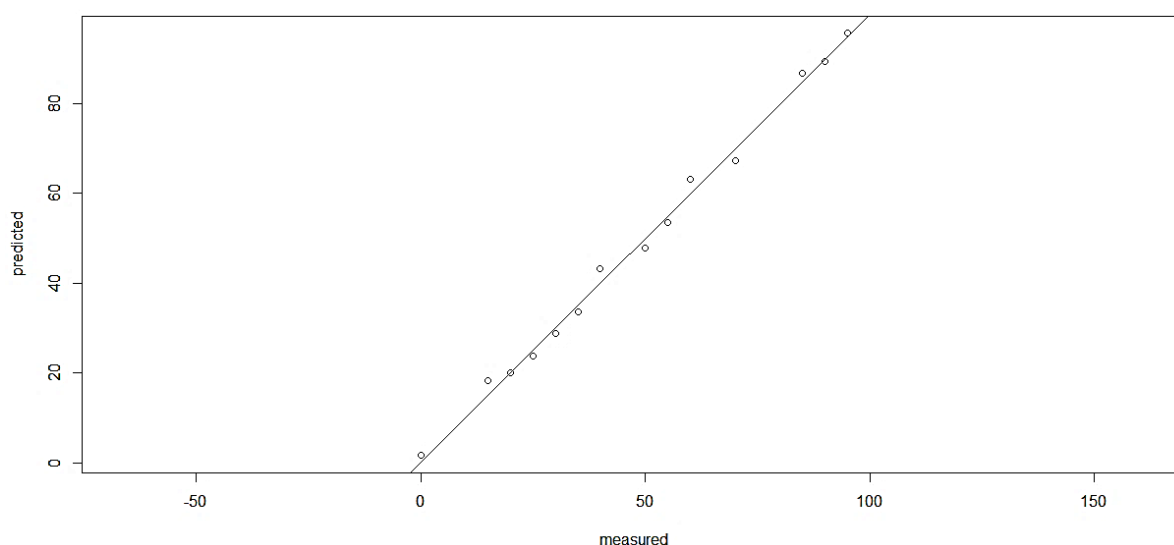
Avaliando os resíduos na tabela 6, após o I-PLS com todas as misturas, verificou-se que as misturas (*blends*): B01, B02, B09, B13, B15, B16 e S01 (100%) ficaram um pouco distante da curva de calibração, portanto foram retirados do modelo de calibração. Apesar das misturas B10 e B16 apresentarem valores maiores de 2,5; elas não inflam o modelo, assim como as misturas B17, B18 e B19 que estão abaixo de -2,5, segundo o intervalo de alavancagem. Na tabela 6, também pode ser visto a comparação da predição antes e depois da validação do modelo I-PLS. A coluna Predição<sup>2</sup> mostra os valores preditos das concentrações das misturas com 14 pontos (misturas que não foram excluídas). Ou seja, dos 21 pontos testados na Predição<sup>1</sup> apenas 14 pontos participaram da validação do modelo.

Tabela 6: Misturas (*blends*) utilizadas na construção do modelo de calibração e validação

Código	Teor experimental	Predição <sup>1</sup> , %	Predição <sup>2</sup> , %
A01	0% de soja	-5,29	0,11
B01	5% de soja	12,80	-
B02	10% de soja	15,82	-
B03	15% de soja	17,02	17,46
B04	20% de soja	20,68	20,06
B05	25% de soja	24,33	24,29
B06	30% de soja	29,07	29,03
B07	35% de soja	34,95	33,89
B08	40% de soja	41,95	41,58
B09	45% de soja	42,39	-
B10	50% de soja	49,47	48,15
B11	55% de soja	51,97	53,89
B12	60% de soja	61,33	62,06
B13	65% de soja	53,85	-
B14	70% de soja	68,11	67,91
B15	75% de soja	80,62	-
B16	80% de soja	75,34	-
B17	85% de soja	89,22	86,36
B18	90% de soja	93,29	89,65
B19	95% de soja	99,84	95,49
S01	100% de Soja	93,21	-

Também foi gerado um novo gráfico da correlação dos valores experimentais e preditos pelo modelo (figura 29). Para o conjunto de validação, obteve-se um valor de RMSEP de 30,58, muito próximo do valor com todas as misturas que foi de 31,79. Os valores experimentais ficaram mais próximos dos valores preditos, portanto pôde ser utilizado para avaliar o teor de adulteração das amostras comerciais com concentração desconhecida. As amostras comerciais também apresentaram boa correlação com o modelo de calibração.

Figura 29: Correlação entre os valores de soja predito (*predict*) pelo modelo e adicionados às misturas (*measured*) após a validação do modelo com 2 componentes.



### 5.2.2.3 Teste com as amostras comerciais de andiroba.

Os resultados obtidos na análise I-PLS foram expostos na tabela 7, em que pode ser identificado o percentual de adulteração no óleo de andiroba por óleo de soja. Como previsto nas análises exploratórias, a amostra A05 foi a que apresentou menor percentual de adulteração, seguido das amostras A15, A10 e A25, que se agrupavam com o padrão de andiroba A06, em que eram classificadas como amostras de adulteração fina. Contudo, apenas a amostra A05 pode ser denominada como adulteração fina, pois as amostras A15, A10 e A25 possuem 23,70-31,86% de adulterante. Quanto a amostra A23 com um percentual de 35,09% ficou entre média a alta adulteração, juntamente com as amostras A15, A10 e A25. As amostras A09 e A24 também continuaram com o mesmo comportamento da análise exploratória, assim como as amostras A26, A22 e A12, que sempre se agrupavam com a soja, cujo grau de adulteração aumenta nessa mesma ordem (44,01-97,01% de soja). As

amostras A26, A22 e A12 apresentam teor de soja em níveis elevados, como se houvesse a adição de andiroba no óleo de soja, que não era a proposta problema do trabalho. Logo, todas as amostras obtidas em feiras da cidade de Manaus foram consideradas adulteradas.

A amostra A12 que se apresentava diferenciada nas análises exploratórias, principalmente no mapa de calor, cuja coloração variava de branco a amarelo, poderia ter sido considerada uma amostra com comportamento anômalo, no entretanto, através do método I-PLS foi possível verificar que o comportamento diferenciado é explicado pelo alto valor percentual de soja nessa amostra. Logo, seria precipitado chamá-la de outlier, pois as características de adulteração possuem níveis muito diferentes entre as amostras comerciais.

Tabela 7: Resultados de concentração (% v/v) de óleo de soja em amostras comerciais pelo método I-PLS

<b>Código</b>	<b>Concentração, %</b>	<b>Predição, %</b>	<b>Classificação</b>
A05	Desconhecido	14,06	Adulterada
A09	Desconhecido	48,37	Adulterada
A10	Desconhecido	31,36	Adulterada
A12	Desconhecido	97,07	Adulterada
A15	Desconhecido	23,70	Adulterada
A22	Desconhecido	71,11	Adulterada
A23	Desconhecido	35,09	Adulterada
A24	Desconhecido	44,01	Adulterada
A25	Desconhecido	31,86	Adulterada
A26	desconhecido	66,91	Adulterada

## 6. CONCLUSÃO

A ressonância magnética nuclear (RMN de  $^1\text{H}$  e bidimensionais) apresentou-se bastante promissora para a caracterização da cadeia graxa dos óleos vegetais utilizados no estudo, pois além de ser uma técnica bastante reprodutível, não houve a necessidade de realizar um tratamento prévio das amostras. O experimento de RMN de  $^1\text{H}$  também é bastante rápido, mesmo se considerarmos os procedimentos de calibração de pulso e estimativa de T1, em média as análises ocorreram num tempo de 34 minutos. Os dados espectrais de RMN de  $^1\text{H}$  quando combinados à análise quimiométrica mostrou-se uma poderosa ferramenta na análise de adulteração do óleo de andiroba com óleo de soja, pois conseguiram diferenciar os perfis espectrais dos óleos de andiroba padrão aos perfis de óleo de andiroba comercial adulterados, bem como determinar as amostras comerciais íntegras.

O método HCA permitiu identificar previamente o comportamento de dois grupos grandes, em que o grupo G1 com as amostras comerciais de andiroba foram consideradas íntegras, devido se agruparem com as amostras padrão; o grupo G2 foram estimados como óleo de andiroba adulterado, pois apresentaram uma distância bastante pronunciada com relação ao grupo G1, mostrando assim a dissimilaridade entre os grupos. Para corroborar essa hipótese, repetiu-se a mesma análise exploratória acrescentando os dados espectrais do óleo de soja, que é bastante utilizado como óleo adulterante de vários óleos vegetais. Logo, também foi observado o mesmo comportamento de agrupamento de G1 e G2, onde as amostras do grupo G2 se agruparam com as amostras de óleo de soja, confirmando a presença de óleo de soja nessas amostras comerciais.

O método FCM conseguiu verificar os mesmos agrupamentos que a HCA, porém foi observada uma divisão intragrupo do grupo G1, que também foi visto anteriormente no mapa de calor e no gráfico de silhueta. Essa subdivisão do grupo G1 foi denominada como G1A e G1B, em o grupo G1A foi formado pelas amostras consideradas íntegras em HCA mais o padrão de andiroba A06, já o grupo G1B foi formado apenas pelas amostras padrão adquiridas do município de Manaquiri. Assim, acredita-se que as amostras foram separadas pela origem em que foram produzidas.

No método PCA robusta além de fazer a compressão dos dados e discriminar as amostras comerciais adulteradas e íntegras, pôde-se identificar as variáveis que pesaram no comportamento entre as amostras. Para as amostras padrão e íntegras

de andiroba, as principais variáveis responsáveis por essa discriminação foram os *buckets* 29, 25 e 18. Esses *buckets* são característicos de cadeia graxa saturada que representa 40,2% do óleo de andiroba, associado aos sinais metilas de cadeias alquila, metileno  $\alpha$  e  $\beta$  carbonila de cadeias alquilas, respectivamente. Portanto, a discriminação dos grupos de andiroba padrão e comercial íntegra se dão principalmente pela presença dos ácidos palmítico e esteárico. Já para as amostras consideradas adulteradas e as amostras de soja, ambas foram discriminadas principalmente pelos *buckets* 28, 22, 14, 4 e 3, sinais característicos de ésteres ou ácidos graxos insaturados. Quando ocorre a adição de óleo de soja no óleo de andiroba, os sinais do óleo de andiroba ficam parecidos com o perfil da soja, ocorrendo um pequeno aumento na área desses sinais. Assim, os ésteres linoleato e linolenato são os responsáveis pela discriminação das amostras adulteradas e amostras de soja, pois estes fazem parte de 59% da cadeia graxa do óleo de soja e não de andiroba. Apenas duas amostras puderam ser denominadas como amostras de andiroba comercial íntegra, que foram as amostras A11 e A14. O método “mad” para PCA robusta foi imprescindível para avaliar o comportamento das amostras adulteradas, uma vez que quando somada as informações do I-PLS, pôde-se entender o porquê do comportamento das amostras na PCA comum; como as amostras comerciais apresentaram variações muito discrepantes entre si, elas se comportavam como *outliers*, o que poderia levar a uma discriminação duvidosa, já que a PCA comum se move na direção de máxima variância.

O método I-PLS foi importante para estimar o grau de adulteração das amostras comerciais de andiroba que foram denominadas como adulteradas nos métodos exploratórios. Assim, o método de calibração possibilitou quantificar o percentual de óleo de soja adicionado nas amostras comerciais de andiroba, em que foi observado cerca de 14,06 a 97,07% de óleo de soja nas amostras comerciais adquiridas nas feiras da cidade de Manaus. Com isso, o modelo aplicado de calibração multivariada I-PLS, apresentou a vantagem da determinação direta, simultânea e rápida do conteúdo de soja em amostras de óleo de andiroba comercial sem a presença de padrões individuais dos ácidos ou ésteres presentes na cadeia graxa dos óleos estudados.

## 7. REFERÊNCIAS

ABBAS, O.; BAETEN, V. Advances in the Identification of Adulterated Vegetable Oils. **Advances in Food Authenticity Testing**, p. 519-542, 2016.

ALMEIDA, M. R. et al. Classification of Amazonian rosewood essential oil by Raman spectroscopy and I-PLS-DA with reliability estimation. **Talanta**, V. 117, p. 305-311, 2013.

ALONSO-SALCES, R. M. et al. Quality Assessment of Olive Oil by <sup>1</sup>H-NMR Fingerprinting. **Food Control**, V. 22, p. 2041-2046, 2011.

AMBROZIN, A. R. P. et al. Limonoids from andiroba oil and *Cedrela fissilis* and their insecticidal activity. **Journal Brazilian Chemical Society**, V. 17, p. 542-547, 2006.

AUBLET, F. N., 1775. *Planta História da Guiana Françoise*.2, Supl. 32. Google Scholar.

BAMBA, T.; FUKUSAKI, E. Technical problems and practical operations in plant metabolomics. **J. Pestic. Sci.**, 31(3), p. 300–304, 2006.

BARISON, A. et al. A simple methodology for the determination of fatty acid composition in edible oils through <sup>1</sup>H NMR spectroscopy. **Magnetic Resonance in Chemistry**, V. 48, p. 642–650, 2010.

BEZDEK, J. C., EHRLICH, R., FULL, W. FCM: The *Fuzzy C-Means* clustering algorithm. **Computers & Geosciences**, V. 10, p. 191-203, 1984.

BRERETON, R.G. "A short history of chemometrics: a personal view". **Journal of Chemometrics**, 28(10): 749, 2014.

CABRAL, E. C. et al. Typification and quality control of the andiroba (*Carapa guianensis*) oil via mass spectrometry fingerprinting. **Analytical Methods**, V. 5, p. 1385-1391, 2013.

CAVALCANTE, F.J.B. et al. Pesquisa e identificação de espécies oleaginosas nativas da Amazônia. **Relatório Técnico**. Convênio CODEAMA/ INPA, Manaus, Amazonas, 1986.

CONAB. Indicadores da Agropecuária - Companhia Nacional de Abastecimento. Brasília, Ano XXVIII, nº 2, p. 01-90, 2019.

CUADROS-RODRÍGUEZ, L. et al. Chromatographic fingerprinting: An innovative approach for food 'identification' and food authentication e A tutorial. **Analytica Chimica Acta**, V. 909, p. 9-23, 2016.

DA SILVA, E. N. et al. Aspectos socioeconômicos da produção extrativista de óleos de andiroba e de copaíba na floresta nacional do Tapajós, Estado do Pará. **Revista de Ciências Agrárias**, V. 53 nº 1, p. 12-23, 2010.

DE SANTANA, F.B. et al. Rapid Discrimination Between Authentic and Adulterated Andiroba Oil Using FTIR-HATR Spectroscopy and Random Forest. **Food Analytical Methods**, V. 11, p.1927-1935, 2018.

DHAKAL, K. H. et al. Variation of unsaturated fatty acids in soybean sprout of high oleic acid accessions. **Food Chemistry**, V. 164, p.70-73, 2014.

FERRAZ, I.D.K.; CAMARGO, J.L.C.; SAMPAIO, P.T.B. Sementes e plântulas de andiroba (*Carapa guianensis* Aubl. e *Carapa procera* D. C.): Aspectos botânicos, ecológicos e tecnológicos. **Acta Amazonica**. V. 32, n. 4, p. 647-661, 2002.

FERREIRA, M. C. Quimiometria – Conceitos, métodos e aplicações. Campinas-SP: Editora Unicamp, 2015.

FLORES, T.B. Meliaceae *in* Flora do Brasil 2020 em construção. **Jardim Botânico do Rio de Janeiro**. Disponível em: <<http://reflora.jbrj.gov.br/reflora/floradobrasil/FB19737>>.

FISCH, S. T. V.; FERRAZ, I. D. K.; RODRIGUES, W. A. Distinguishing *Carapa Guianensis* Aubl. from *Carapa procera* D.C. (Meliaceae) by morphology of young seedlings. **Acta Amazônica**, V. 25, p. 193-200, 1995.

FUNARI, C. S. et al. Metabolômica, uma abordagem otimizada para exploração da biodiversidade brasileira: estado da arte, perspectivas e desafios. **Química Nova**, V. 36, Nº 10, p. 1605-1609, 2013.

GASPAR AMARAL, L. F.; FIERRO, I. M. Profile of medicinal plants utilization through patent documents: the andiroba example. **Revista Brasileira de Farmacognosia**, V. 23, n.4, p. 716-722, 2013.

GUEDES, M.C. et al. Produção de sementes e óleo de andiroba (*Carapa guianensis* Aubl.) em área de várzea do Amapá. *In: Seminário do Projeto Kamukaia Manejo Sustentável de Produtos Florestais Não Madeireiros na Amazônia*. Embrapa Acre, 2008.

GÓMEZ-CARAVACA, A. M.; MAGGIO, R. M.; CERRETANI, L. Chemometric applications to assess quality and critical parameters of virgin and extra-virgin olive oil. A review. **Analytica Chimica Acta**, V. 913, p. 1-21, 2016.

INDELICATO, S. et al. Triacylglycerols in edible oils: Determination, characterization, quantitation, chemometric approach and evaluation of adulterations. **Journal of Chromatography A**, 1515, p. 1–16, 2017.

KAUFMAN, L.; ROUSSEUW, P. J. Finding Groups in Data: An Introduction to Cluster Analysis. **Wiley**: New York, 2005.

KENFACK, D. Resurrection in *Carapa* (Meliaceae): a reassessment of morphological variation and species boundaries using multivariate methods in a phylogenetic context. **Botanical Journal of the Linnean Society**, V. 165, p. 186–221, 2011(a).

KENFACK, D. A Synoptic Revision of *Carapa* (Meliaceae). **Harvard Papers in Botany**. V. 16, n. 2, p. 171-231, 2011(b).

KORIFI, R. et al. Composition and authentication of virgin olive oil from French PDO regions by chemometric treatment of Raman spectra. **Journal of Raman Spectroscopy**, V. 42, n. 7, p. 1540-1547, 2011.



KOOY, F. V. D. et al. Quality Control of Herbal Material and Phytopharmaceuticals with MS and NMR Based Metabolic Fingerprinting. **Planta Medica**, V. 75, p. 763–775, 2009.

LAMANNA, R. Proton NMR Profiling of Food Samples (Ed: G. A. Webb), **Annual Reports on NMR Spectroscopy**, V. 80, Oxford, United King-dom, p. 239–291, 2013.

LIRA-GUEDES, A. C.; NARDI, M. Guia prático para o manejo sustentável de andirobeiras de várzea e para a extração do óleo de suas sementes. **EMBRAPA**, ISBN: 978-85-7035-500-3, 2015.

Missouri Botanical Garden. <http://www.tropicos.org/Name/20400362>. >. Acesso em: 13 setembro de 2018.

MAPA - Ministério da Agricultura, Pecuária e Abastecimento. PORTARIA Nº 14, DE 3 DE JANEIRO DE 2018. **Diário Oficial da União**, Nº 9, sexta-feira, 12 de janeiro de 2018.

MEZIANE-ASSAMI, D. et al. Geographical Differentiation of Rosemary Based on GC/MS and Fast HPLC Analyses. **Food Analytical Methods**, V.6, p. 282-288, 2013.

MIRANDA JR, R. N. C. et al. Antiplasmodial activity of the andiroba (*Carapa guianensis* Aubl., Meliaceae) oil and its limonoid-rich fraction. **Journal of Ethnopharmacology**, V.142, Issue 3, p. 679-683, 2012.

MONFREDA, M.; GOBBI L. A. Grippa. *Blends* of olive oil and sunflower oil: characterisation and olive oil quantification using fatty acid composition and chemometric tools. **Food Chemistry**, V. 134, p. 2283-2290, 2012.

GOWDA, G. A. N.; RAFTERY, D. Advances in NMR-based metabolomics. **Comprehensive Analytical Chemistry**. V. 63, p. 187-211, 2014.

OBISESAN, K. A. et al. HPLC-UV and HPLC-CAD chromatographic data fusion for the authentication of the geographical origin of palm oil. **Talanta**, V. 170, p. 413-418, 2017.

PENNINGTON, T. D.; STULES, B. T.; TAYLOR, D. A. H. Meliaceae. **Flora Neotrópica**, V. 28, p. 406-419, 1981.

PIZARRO, C. et al. Recognition of volatile compounds as markers in geographical discrimination of Spanish extra virgin olive oils by chemometric analysis of nonspecific chromatography volatile profiles. **Journal Chromatography A**, V. 1218, p. 518-523, 2011.

POPESCU, R. et al. Discrimination of vegetable oils using NMR spectroscopy and chemometrics. **Food Control**, V. 48, p. 84-90, 2015.

REZANKA, T; PADROVA, K.; SIGLER, K. Regioisomeric and enantiomeric analysis of triacylglycerols. **Analytical Biochemistry**, V. 524, p. 3-12, 2017.

RUIZ-SAMBLÁS, C. et al. Quantification of blending of olive oils and edible vegetable oils by triacylglycerol fingerprint gas chromatography and chemometric tools. **Journal Chromatography B**, V. 910, p. 71e77, 2012.

SANTOS, J. S. et al. <sup>1</sup>H NMR combined with chemometrics tools for rapid characterization of edible oils and their biological properties. **Industrial Crops and Products**, V. 116, p. 191-200, 2018.

SAVORANI, F; TOMASI, G; ENGELSEN, S. B. ICOSHIFT: A versatile tool for the rapid alignment of 1D NMR spectra. **Journal of Magnetic Resonance**, V. 202, p. 190–202, 2010.

SENA, M. M. e POPPI, R. J. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, 23 (4), p. 547-556, 2000.

SHANLEY, P e MEDINA, G. Andiroba (*Carapa guianensis*, Aublet.). Frutíferas e plantas úteis na vida amazônica. Belém: Cifor, 2005, p.41-50.

SHI, T. et al. <sup>1</sup>H NMR combined with chemometrics for the rapid detection of adulteration in camellia oils. **Food Chemistry**. V. 242, p. 308-315, 2018.

SINGH, K.; BLÜMICH, B. NMR spectroscopy with compact instruments. **Trends in Analytical Chemistry**, V.83, p. 12–26, 2016.

SILVA, S. G; NUNOMURA, R. C. S; NUNOMURA, S. M. Limonoides isolados dos frutos de *Carapa guianensis* Aublet (Meliaceae). **Química Nova**, V. 35, Nº 10, p. 1936-1939, 2012.

SINELLI, N. et al. Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity. **Food Research International**, V. 43, 369-375, 2010

SOMPILA, A. W. G. T. et al. Fast non-aqueous reversed-phase liquid chromatography separation of triacylglycerol regioisomers with isocratic mobile phase. Application to different oils and fats. **Journal of Chromatography B**, V. 1041–1042, p. 151-157, 2017.

TSIMIDOU, M.Z. et al. Food Fraud. **Encyclopedia of Food and Health, Academic Press**, p. 35-42, 2016.

TYAGI, S. et al. Applications of metabolomics - A systematic study of the unique chemical fingerprints: an overview. **Int. J. Pharm. Sci. Rev. Res**, V. 3, nº 1, p. 83-86, 2010.

VERPOORTE, R. et al. NMR-based metabolomics at work in phytochemistry. **Phytochemistry Reviews**, V. 6, p. 3–14, 2007.

VERPOORTE, R. et al. Metabolomics: back to basics. **Phytochemistry Reviews**, V. 7, p. 525–537, 2008.

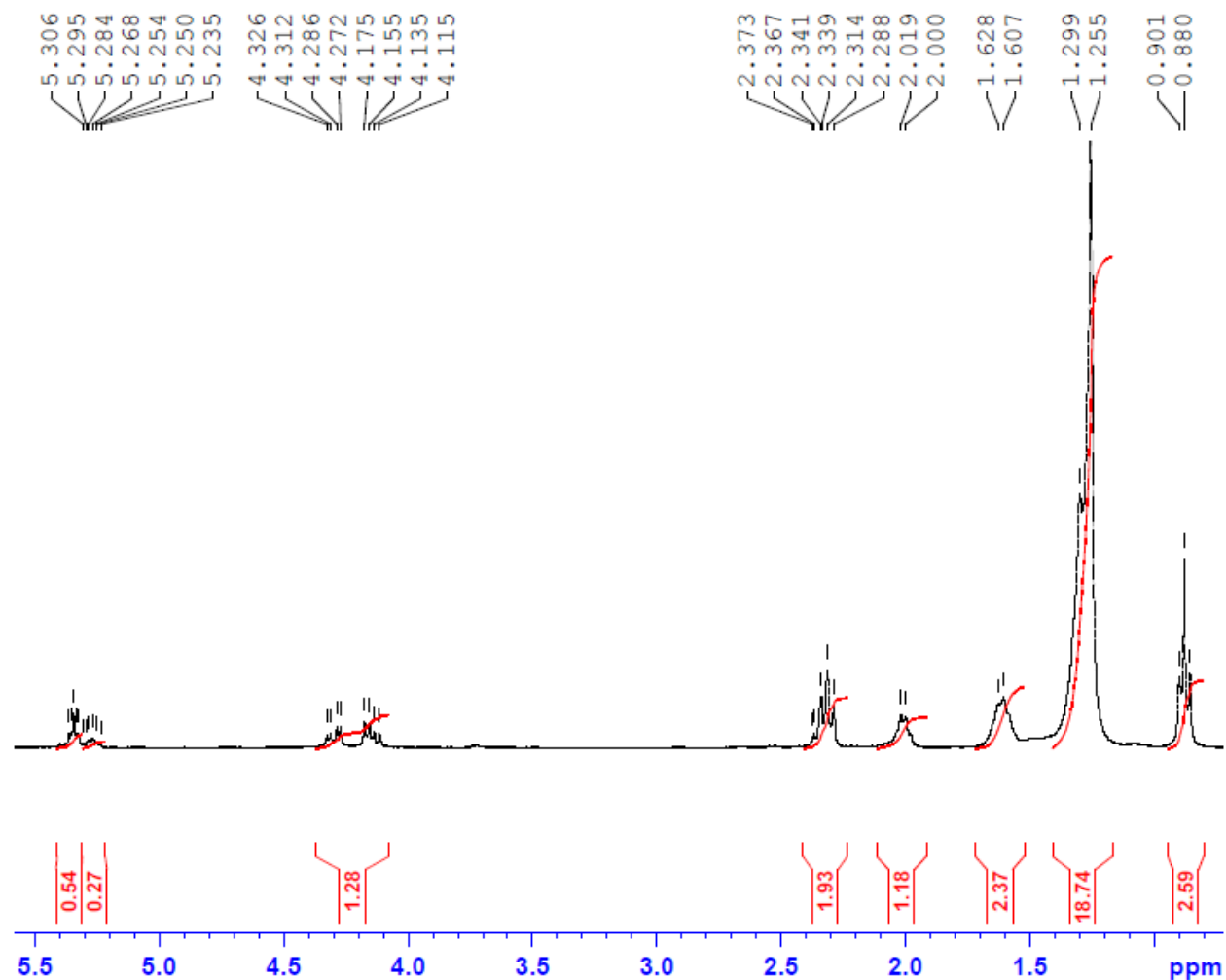
VIEGAS JR, C.; BOLZANI, V.S. Os Produtos Naturais e a Química Medicinal Moderna. **Química Nova**, 29(2), p. 326-337, 2006.

WOLD, S. "Chemometrics; what do we mean with it, and what do we want from it?". **Chemometrics and Intelligent Laboratory Systems**, 30(1): 109, 1995.

ZHU, W.; WANG, X.; CHEN, L. Rapid detection of peanut oil adulteration using low-field nuclear magnetic resonance and chemometrics. **Food Chemistry**, V. 216, p. 268-274, 2017.

## 8. ANEXOS

Berna18\_And1 (1H, 50uL de oleo + 500 uL CDC13)  
 Op.: Magno



```

Current Data Parameters
NAME      Berna18_And1
EXPNO    1
PROCNO   1

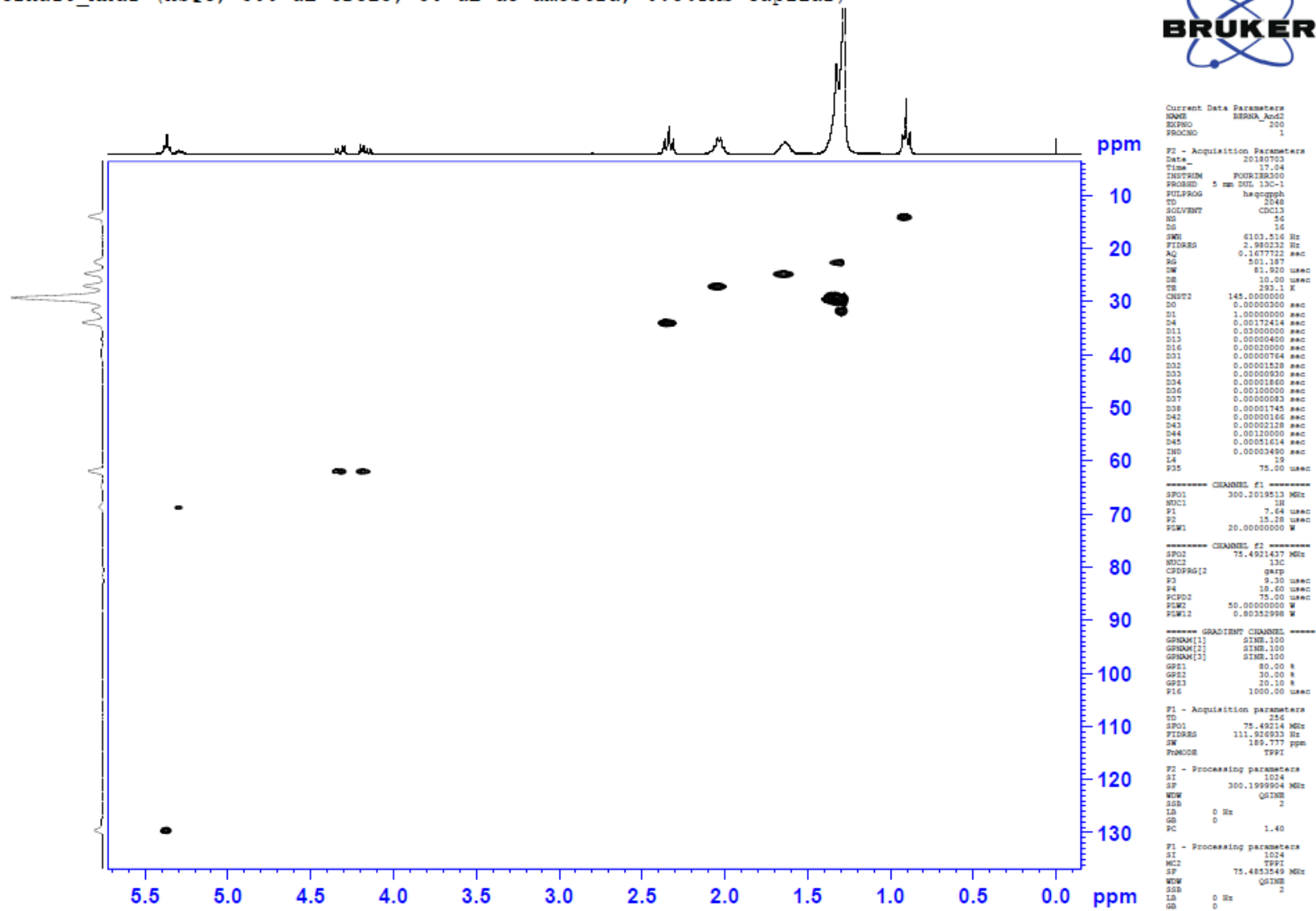
F2 - Acquisition Parameters
Date_    20180904
Time_    17.21
INSTRUM  FOURIER300
PROBHD   5 mm DUL 13C-1
PULPROG  zg
TD        65536
SOLVENT  CDC13
NS        32
DS         2
SWH       6103.516 Hz
FIDRES    0.093132 Hz
AQ        5.3687091 sec
RG        1.65139
DW        81.920 usec
DE        10.00 usec
TE        1024.1 K
D1        1.00000000 sec
TD0       1

===== CHANNEL f1 =====
SF01     300.2019513 MHz
NUC1      1H
P1        7.88 usec
PLW1     20.00000000 W

F2 - Processing parameters
SI        65536
SF        300.1999977 MHz
WDW       EM
SSB       0
LB        0 Hz
GB        0
PC        1.00
  
```

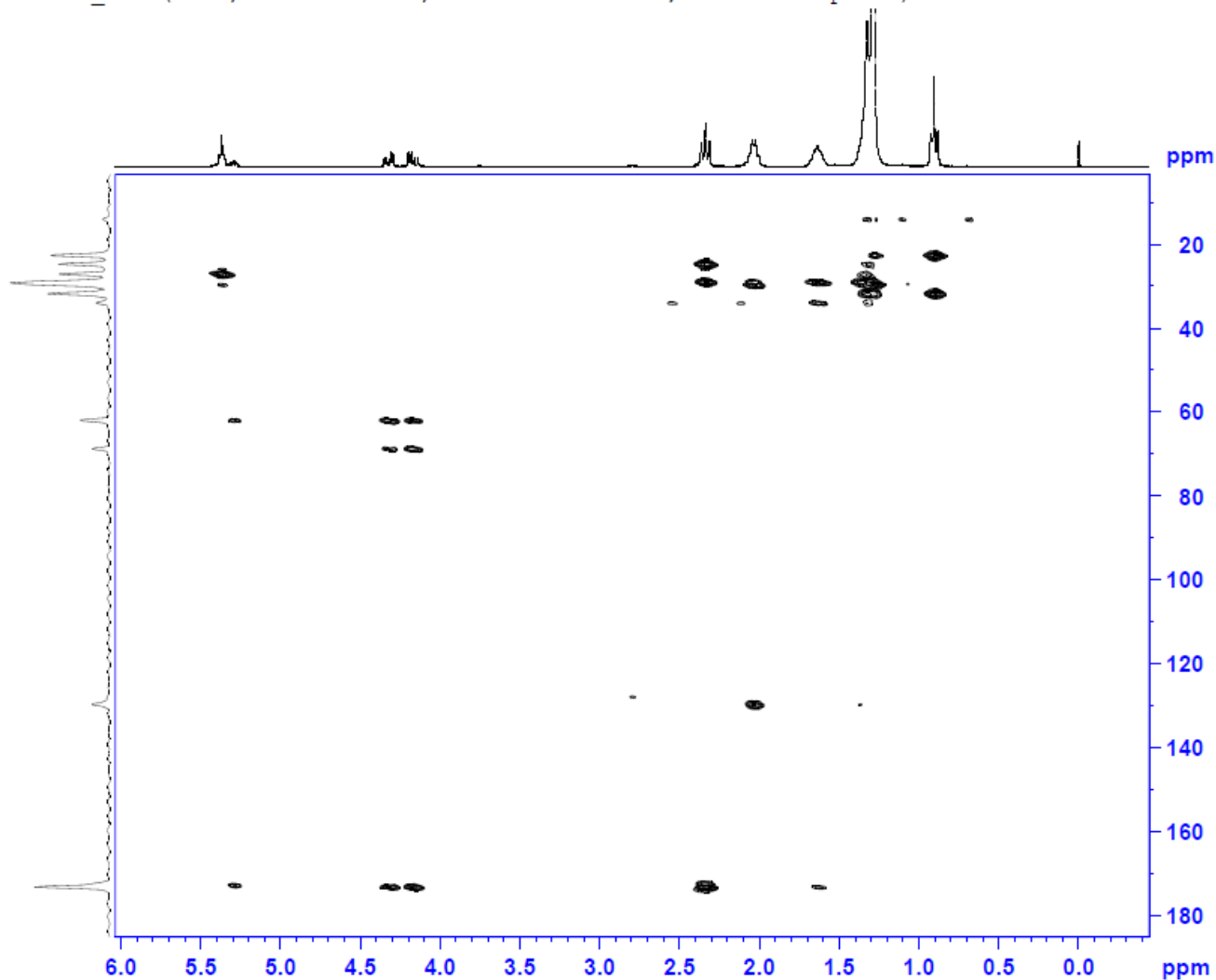
ANEXO A - ESPECTRO DE RMN DE <sup>1</sup>H DO ÓLEO DE ANDIROBA PADRÃO

Berna18\_And1 (HSQC; 500 uL CDCl3; 50 uL de amostra, 0.5%TMS capilar)



ANEXO B – MAPA DE CONTORNO HSQC DO ÓLEO DE ANDIROBA

Berna18\_And1(HMBC; 500 uL CDC13; 50 uL de amostra, 0.5%TMS capilar)



```

Current Data Parameters
NAME      BERNA_And2
EXPNO     300
PROCNO    1

F2 - Acquisition Parameters
Date_     20180703
Time      21.50
INSTRUM   PULPROG
PROBHD    5 mm BBO 13C-1
PULPROG   hmcpgp1pdqf
TD         2048
SOLVENT   CDCl3
NS         112
DS         16
SWH        6103.516 Hz
FIDRES     2.980232 Hz
AQ         0.1677722 sec
RG         501.187
RW         81.920 usec
DE         10.00 usec
TE         297.4 K
CMT2      145.000000
CMT13     7.000000
D0         0.0000300 sec
D1         1.0000000 sec
D2         0.00344828 sec
D6         0.07142857 sec
D16        0.00020000 sec
D31        0.00000764 sec
D32        0.00001528 sec
D33        0.00000930 sec
D36        0.00100000 sec
D38        0.00003010 sec
IN0        0.00003010 sec

----- CHANNEL f1 -----
SF01      300.2019513 MHz
NUC1       13C
P1         7.64 usec
P2         15.28 usec
PLW1      20.00000000 W

----- CHANNEL f2 -----
SF02      75.4932760 MHz
NUC2       13C
P3         9.30 usec
PLW2      50.00000000 W

----- GRADIENT CHANNEL -----
GPNAM[1]   SINE.100
GPNAM[2]   SINE.100
GPNAM[3]   SINE.100
GP21       50.00 %
GP22       30.00 %
GP23       40.10 %
P16        1000.00 usec

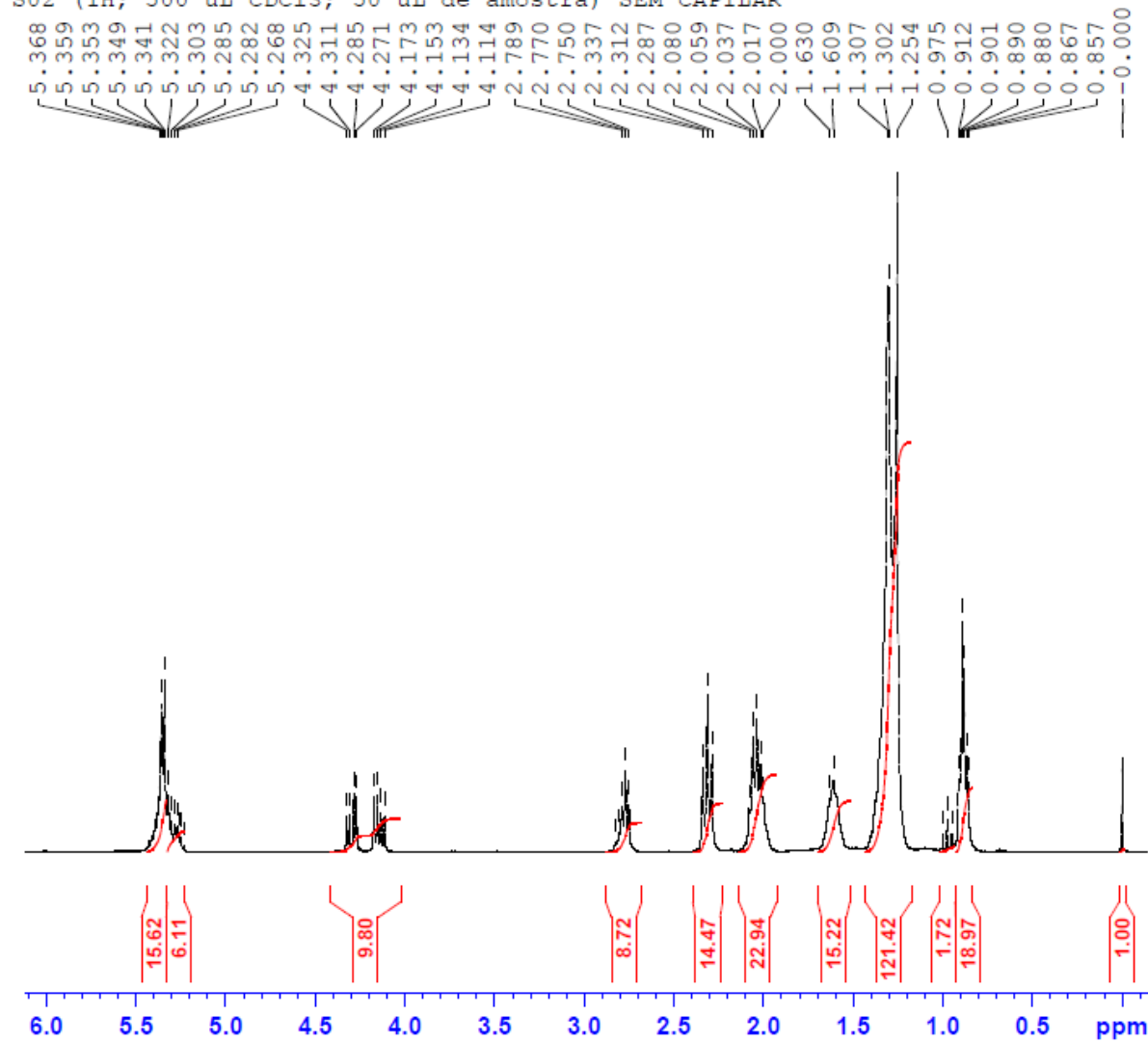
F1 - Acquisition parameters
TD         256
SF01       75.49328 MHz
FIDRES     129.795742 Hz
SW         220.037 ppm
PhMODE     QF

F2 - Processing parameters
SI         1024
SF         300.1999895 MHz
WDW        SINE
SSB        0
LB         0 Hz
GB         0
PC         1.40

F1 - Processing parameters
SI         1024
MC2        QF
SF         75.4853516 MHz
WDW        SINE
SSB        0
LB         0 Hz
GB         0
    
```

ANEXO C - MAPA DE CONTORNO HMBC DO ÓELO DE ANDIROBA

S02 (1H; 500 uL CDC13; 50 uL de amostra) SEM CAPILAR



Current Data Parameters  
 NAME BERN\_A\_Korin  
 EXPNO 11  
 PROCNO 1

F2 - Acquisition Parameters  
 Date\_ 20180711  
 Time\_ 17.49  
 INSTRUM FOURIER300  
 PROBHD 5 mm DUL 13C-1  
 PULPROG zg  
 TD 65536  
 SOLVENT CDC13  
 NS 32  
 DS 2  
 SWH 6103.516 Hz  
 FIDRES 0.093132 Hz  
 AQ 5.3687091 sec  
 RG 1  
 DW 81.920 usec  
 DE 10.00 usec  
 TE 292.6 K  
 D1 1.00000000 sec  
 TDO 1

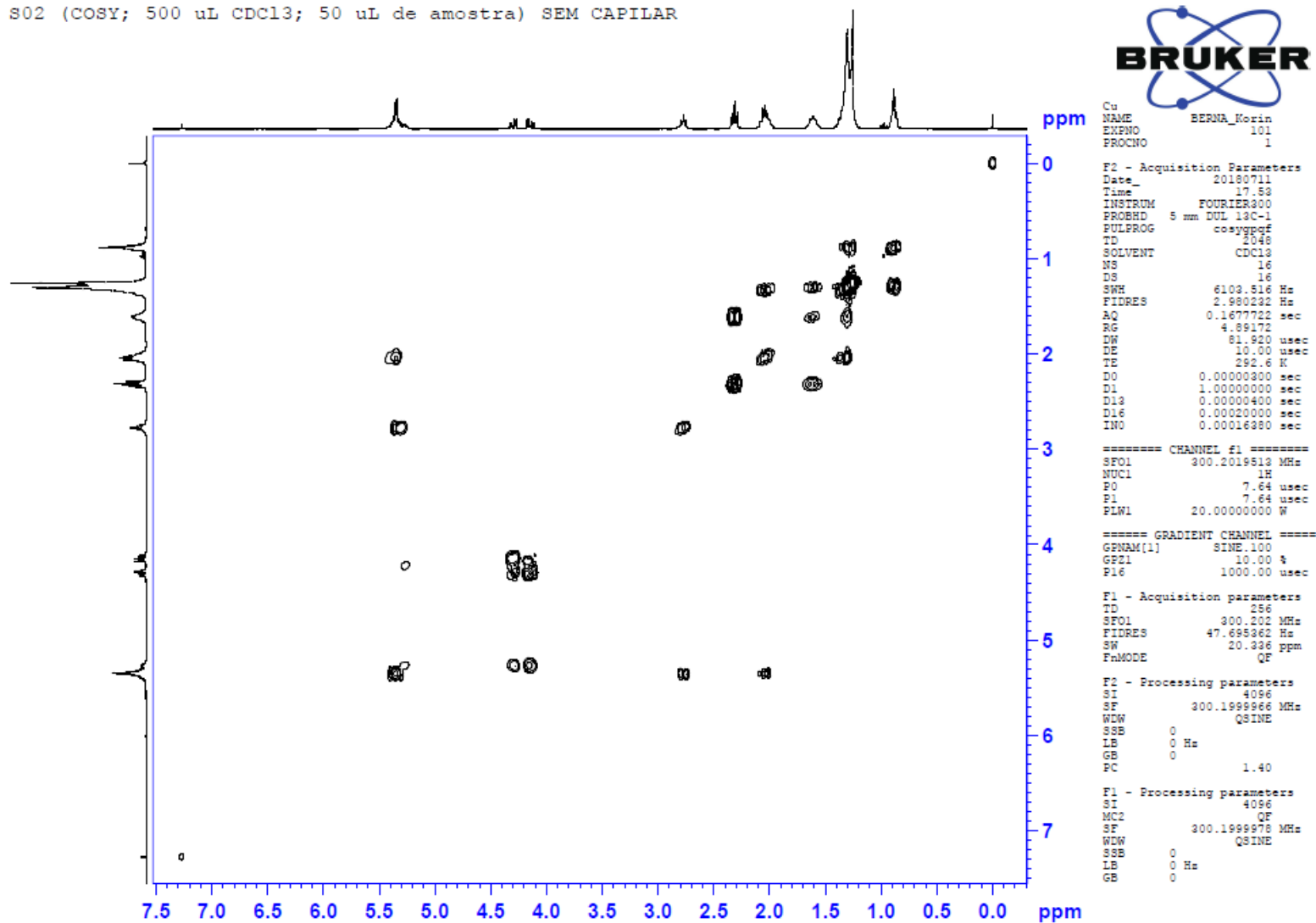
===== CHANNEL f1 =====  
 SFO1 300.2019513 MHz  
 NUC1 1H  
 P1 7.88 usec  
 PLW1 20.00000000 W

F2 - Processing parameters  
 SI 65536  
 SF 300.1999990 MHz  
 WDW EM  
 SSB 0  
 LB 0 Hz  
 GB 0  
 PC 1.00

ANEXO D - ESPECTRO DE RMN DE 1H DO ÓLEO DE SOJA

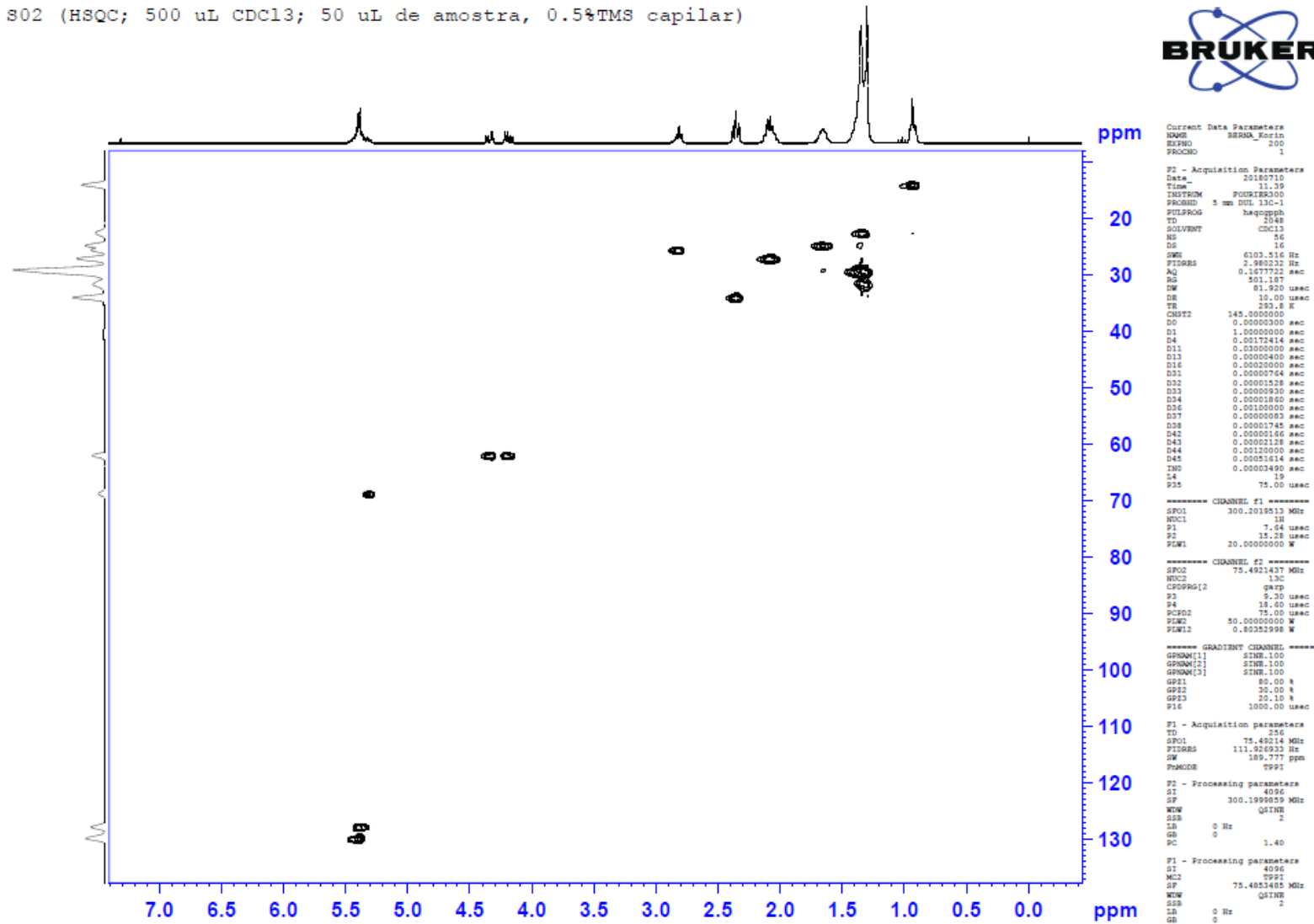


S02 (COSY; 500 uL CDCl3; 50 uL de amostra) SEM CAPILAR



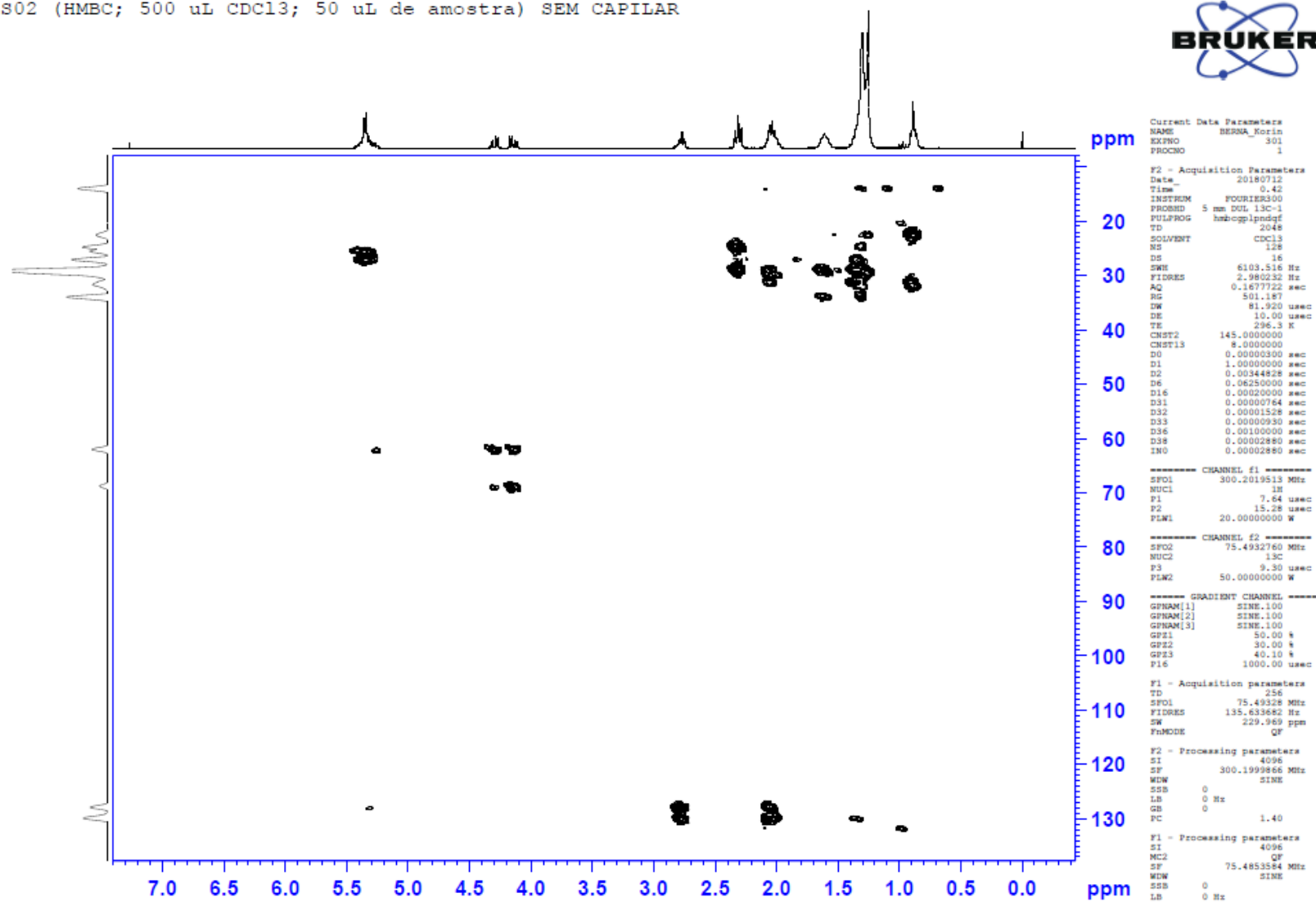
ANEXO E - MAPA DE CONTORNOS COSY DO ÓLEO DE SOJA

S02 (HSQC; 500 uL CDCl3; 50 uL de amostra, 0.5%TMS capilar)

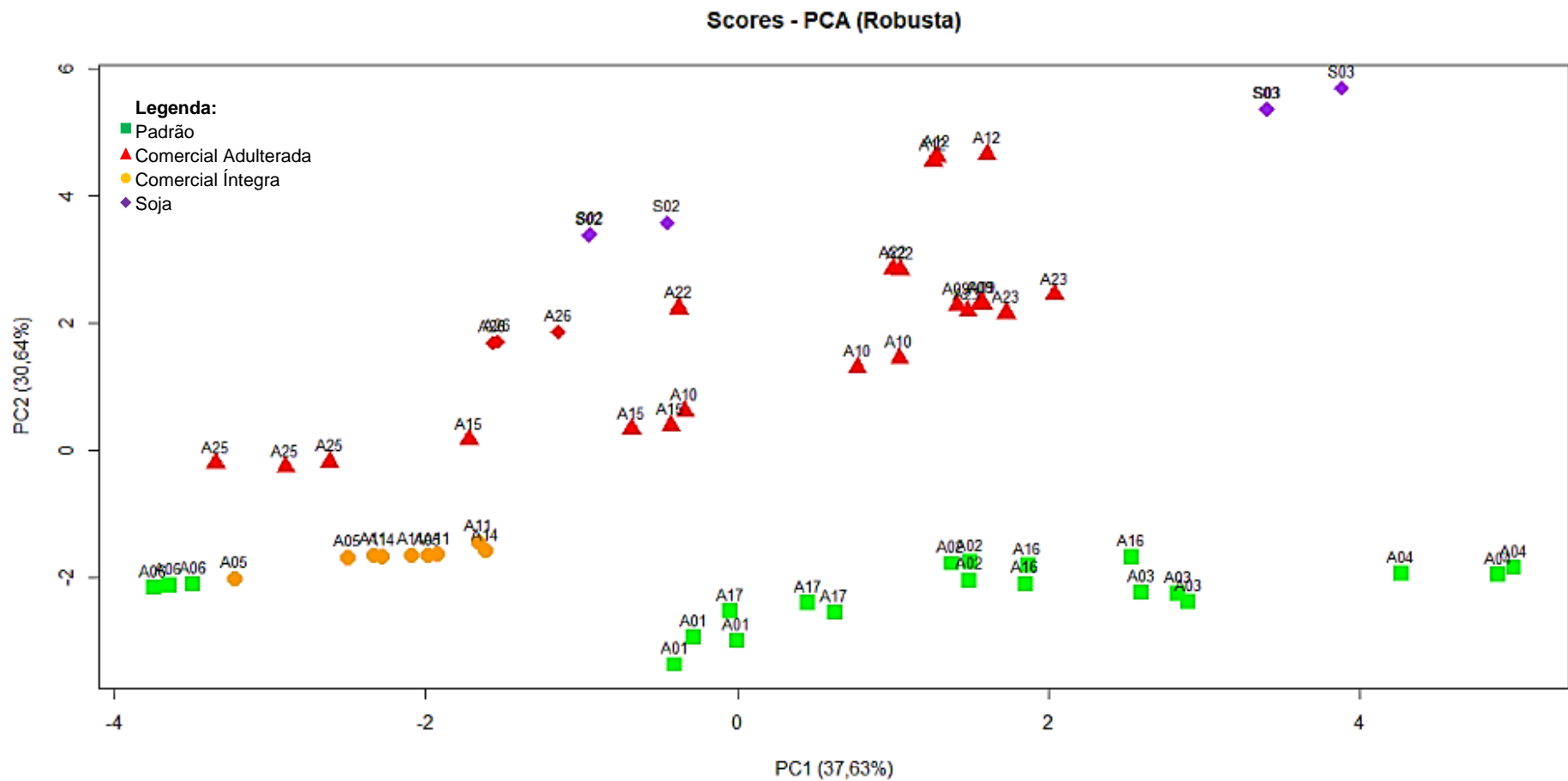


ANEXO F - MAPA DE CONTORNO HSQC DO ÓLEO DE SOJA

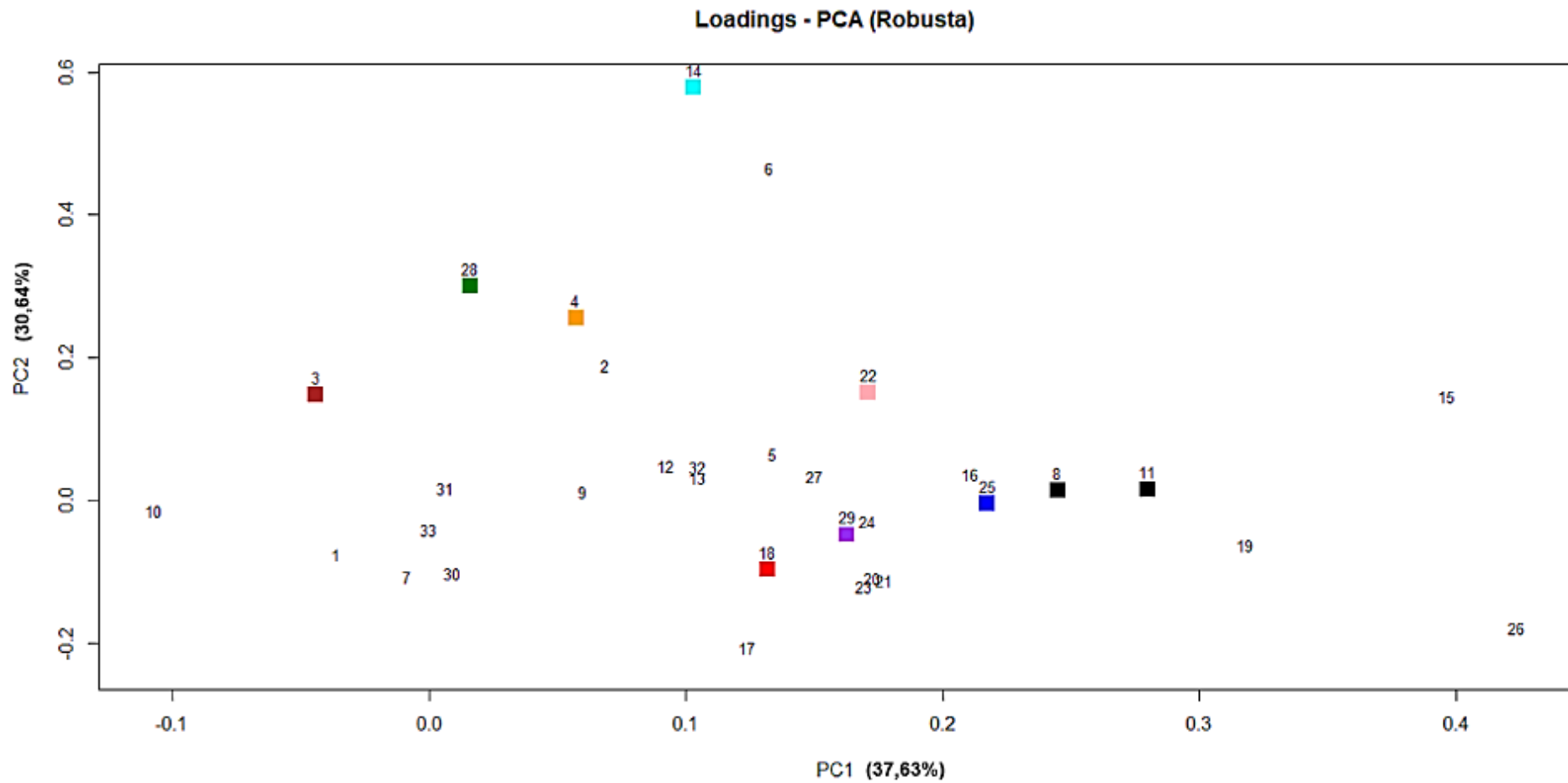
S02 (HMBC; 500 uL CDCl3; 50 uL de amostra) SEM CAPILAR



ANEXO G - MAPA DE CONTORNO HMBC DO ÓLEO DE SOJA

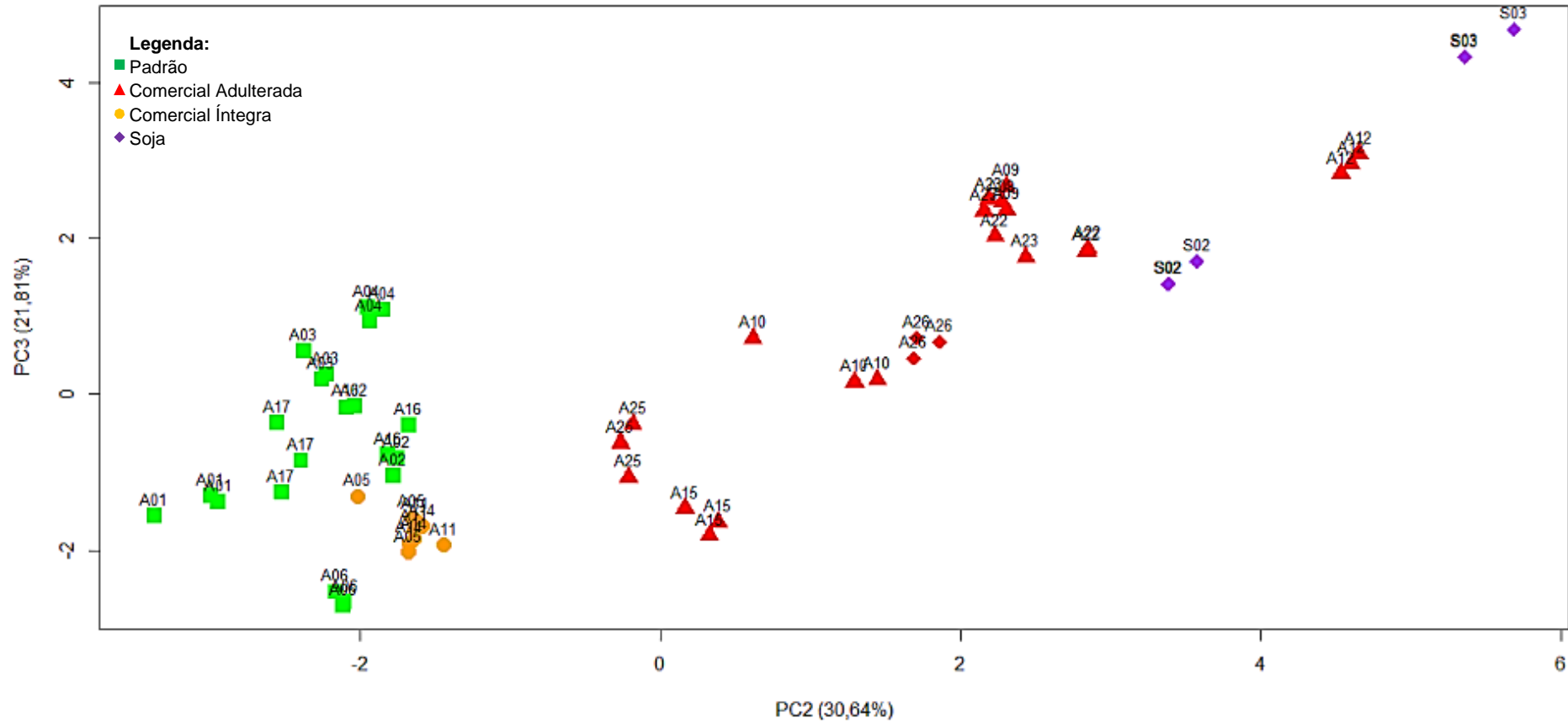


ANEXO H - AMPLIAÇÃO DO GRÁFICO DE SCORES PC1 (37,63%) VERSUS PC2 (30,64%) DAS AMOSTRAS DE ÓLEO DE ANDIROBA

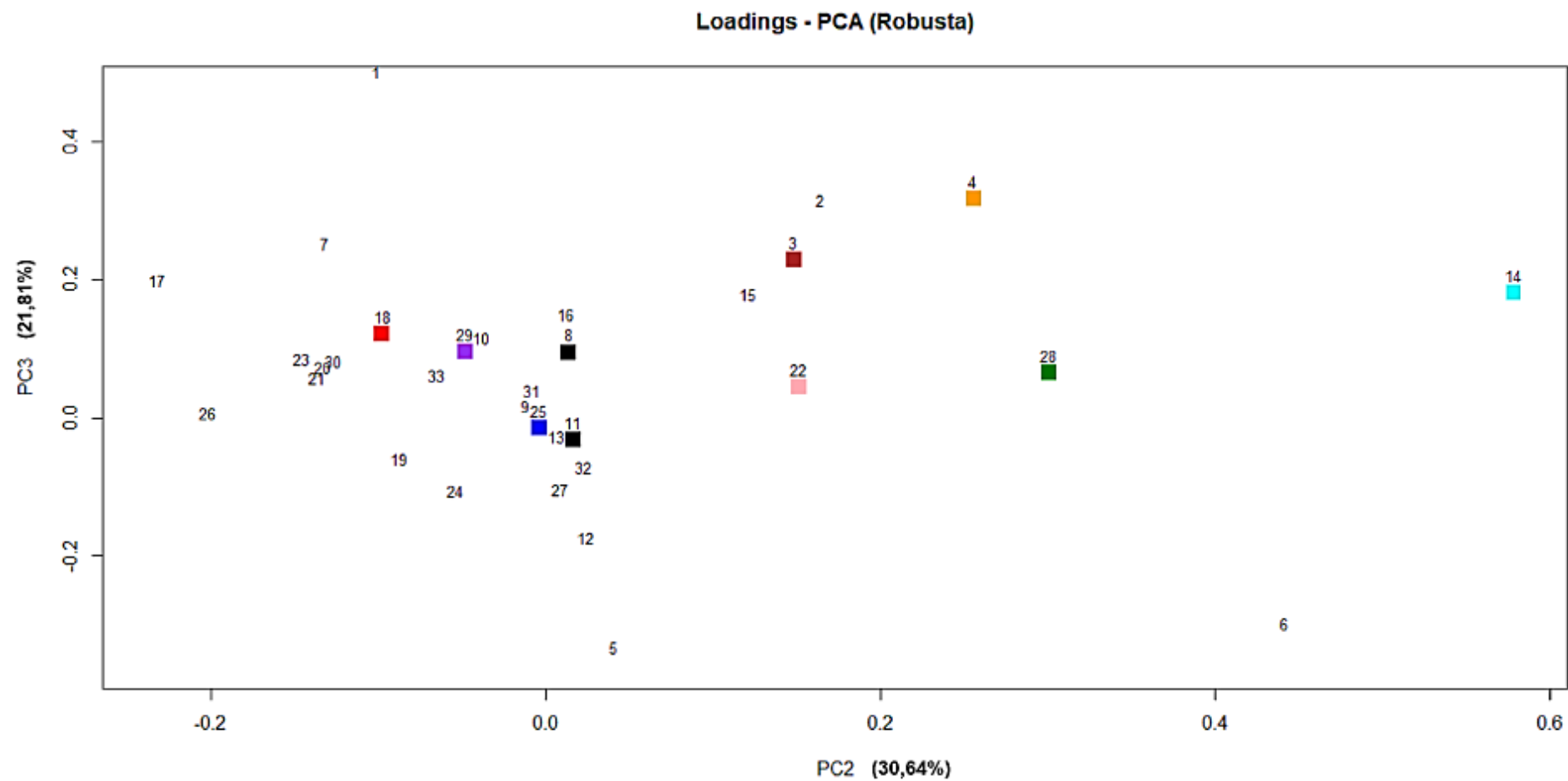


ANEXO I - AMPLIAÇÃO DO GRÁFICO DE *LOADINGS* PC1 (37,63%) *VERSUS* PC2 (30,64%) DESTACANDO AS PRINCIPAIS VARIÁVEIS RESPONSÁVEIS NA DISCRIMINAÇÃO DOS GRUPOS REPRESENTADOS NO ANEXO H

### Scores - PCA (Robusta)



ANEXO J – AMPLIAÇÃO DO GRÁFICO DE SCORES PC2 (30,64%) VERSUS PC3 (21,81%) DAS AMOSTRAS DE ÓLEO DE ANDIROBA PADRÃO, ANDIROBA COMERCIAL ADULTERADA, ANDIROBA COMERCIAL ÍNTEGRA E SOJA



ANEXO K – AMPLIAÇÃO DO GRÁFICO DE *LOADINGS* PC2 (30,64%) *VERSUS* PC3 (21,81%) DESTACANDO AS PRINCIPAIS VARIÁVEIS RESPONSÁVEIS NA DISCRIMINAÇÃO DOS GRUPOS REPRESENTADOS NO ANEXO J.